

GenCore version 5.1.6
Copyright. (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 17:47:37 ; Search time 39 Seconds

(without alignments)
54.262 Million cell updates/sec

Title: US-09-509-482d-21

Perfect score: 118

Sequence: 1 MEWENKRSDWLSVLR TAGVE 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51	43.2	960	2 JH0381	phosphoenolpyruvat
2	51	43.2	960	2 S31159	phosphoenolpyruvat
3	50	42.4	565	1 Q7550	methyalmalonyl-CoA
4	48	40.7	966	1 QVIX1	phosphoenolpyruvat
5	47	39.8	818	2 P9772	hypothetical prote
6	46.5	39.4	323	2 P69373	L-carnitine dehydr
7	46.5	39.4	416	2 P70593	hypothetical prote
8	46.5	39.4	638	2 AB0446	probable exported
9	46	39.0	321	1 F87102	probable phosphoes
10	46	39.0	324	1 B70885	probable phosphoes
11	46	39.0	1168	2 T30935	reverse transcript
12	45	38.1	117	2 T39788	very hypothetical
13	45	38.1	188	2 A13391	hypothetical cytos
14	44.5	37.7	132	2 S22568	hypothetical prote
15	44.5	37.7	132	2 S16058	hypothetical prote
16	44.5	37.7	326	2 S54042	hypothetical prote
17	44.5	37.7	857	2 T20318	hypothetical prote
18	44	37.3	325	2 T05356	hypothetical prote
19	44	37.3	396	2 H72224	hypothetical prote
20	44	37.3	415	2 T32030	hypothetical prote
21	44	37.3	523	2 H90529	potassium uptake p
22	44	37.3	556	2 G69327	DNA ligase (lig) h
23	44	37.3	637	2 A82301	beta-N-acetylhexos
24	44	37.3	715	2 B75135	phosphoenolpyruvat
25	44	37.3	970	1 QYZM	conserved hypotet
26	43.5	36.9	209	2 AF1515	hypothetical prote
27	43.5	36.9	265	2 C83941	hypothetical prote
28	43.5	36.9	436	2 F90790	hypothetical membr
29	43.5	36.9	436	2 AB5651	hypothetical prote

probable transcrip
trNA intron endonu
ABC transporter, p
probable ferric en
hypothetical prote
probable periplasm
conserved hypotet
probable membrane
phenylalanine-tRNA
probable membrane
phosphoenolpyruvat
phosphoenolpyruvat
phosphoenolpyruvat
phosphoenolpyruvat
phosphoenolpyruvat
phosphoenolpyruvat

ALIGNMENTS

RESULT 1

JH0381

phosphoenolpyruvate carboxylase (EC 4.1.1.31) - sorghum

C;Species: Sorghum bicolor

C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 18-Jul-2001

C;Accession: JH0381; S16455

R;Esterlin, C.; Gatti, S.; Keryer, E.; Lepiniec, L.; Tagu, D.; Vidal, J.; Gadal, P.

Gene 99, 87-94, 1991

A;Title: The phosphoenolpyruvate carboxylase gene family of Sorghum: promoter structure

A;Reference number: JH0381; PMID:91216449; PMID:2022326

A;Accession: JH0381

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-960 <CRE>

A;Cross-references: GB:X55664

A;Superfamily: phosphoenolpyruvate carboxylase

C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 43.2%; Score 51; DB 2; Length 960;

Best Local Similarity 56.2%; Pred. No. 14;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWENKRSDWLSVLR 17

Db 471 EWSEKRDWLSVLR 486

RESULT 2

S31159

phosphoenolpyruvate carboxylase (EC 4.1.1.31) CP21 - sorghum

C;Species: Sorghum bicolor (sorghum)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C;Accession: S31159

R;Lepiniec, L.; Keryer, E.; Philippe, H.; Gadal, P.; Cretin, C.

Plant Mol. Biol. 21:487-502, 1993

A;Title: Sorghum phosphoenolpyruvate carboxylase gene family: structure, function and

A;Reference number: S31159; PMID:93184205; PMID:8443342

A;Accession: S31159

A;Molecule type: DNA

A;Residues: 1-960 <LEP>

A;Cross-references: EMBL:X65137; NID:G22614; PIDN:CAA46267.1; PID:G22615

C;Genetics:

A;Introns: 52/3; 183/3; 211/3; 285/3; 320/3; 349/3; 401/3; 734/3; 863/3

C;Superfamily: phosphoenolpyruvate carboxylase

C;Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoprotein

Query Match 43.2%; Score 51; DB 2; Length 960;

Best Local Similarity 56.2%; Pred. No. 14;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWENKRSDWLSVLR 17

|| :||| |||

Db 470 EWSEERQDWLLSEIR 495

RESULT 3

C72550

methyilmalonyl-CoA mutase (EC 5.4.99.2) chain A [similarity] - Aeropyrum pernix (strain R N); Alternate names: protein APE1687

C:Species: Aeropyrum pernix

C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 16-Jun-2000

C:Accession: C72550

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takaiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res: 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72550

A:Molecule type: DNA

A:Residues: 1-565 <RAW>

A:Cross-references: DDBJ:AP000062; NID:G5105244; PIDN:BAA80688.1; PID:G5105375

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1687

C:Function:

A:Description: catalyzes the isomerisation of L-methylmalonyl-CoA to succinyl-CoA

A:Pathway: valine, isoleucine, threonine, methionine, odd-chain fatty acid and cholesterol

C:Superfamily: Streptomyces isobutyryl-CoA mutase chain A

C:Keywords: intramolecular transferase; isomerase

Query Match 42.4%; Score 50; DB 1; Length 565;

Best Local Similarity 54.5%; Pred. No. 12;

Matches 12; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 EW--RNKKRSDWLSWLRTAGV 21

Db 316 EWGAKKRSWLRFRHTOTAGV 337

RESULT 4

QYIX1

phosphoenolpyruvate carboxylase (EC 4.1.1.31) 1 - common ice plant
C:Species: Mesembryanthemum crystallinum (common ice plant)
C:Date: 30-Sep-1991 #sequence/revision 30-Sep-1991 #text_change 18-Jun-1999
C:Accession: S05506; S02716; S26236
R:Cushman, J.C.; Bohnert, H.J.
Nucleic Acids Res 17, 6745, 1989

A:Title: Nucleotide sequence of the gene encoding a CAM specific isoform of phosphoenolpyruvate carboxylase from Mesembryanthemum crystallinum (common ice plant)
A:Reference number: S05506; MUID:89386017; PMID:2780306
A:Accession: S05506
A:Molecule type: DNA

A:Residues: 1-966 <CUS>
A:Cross-references: EMBL:X14587; MID:G22558; PIDN:CAA32727.1; PID:G22559
R:Rickers, J.; Cushman, J.C.; Michalowski, C.B.; Schmitt, J.M.; Bohnert, H.J.
Mol. Gen. Genet. 215, 447-454, 1989

A:Title: Expression of the CAM-form of phospho(enol)pyruvate carboxylase and nucleotide sequence of the gene encoding the CAM-form of phosphoenolpyruvate carboxylase from Mesembryanthemum crystallinum (common ice plant)
A:Reference number: S02716; MUID:89218954; PMID:2710107
A:Accession: S02716
A:Molecule type: mRNA

A:Residues: 1-966 <RIC>
A:Cross-references: EMBL:X13660; MID:G19535; PIDN:CAA31956.1; PID:G19536
R:Cushman, J.C.; Bohnert, H.J.
Plant Mol. Biol. 20, 411-424, 1992

A:Title: Salt stress alters A7'-rich DNA-binding factor interactions within the phosphoenolpyruvate carboxylase gene of Mesembryanthemum crystallinum (common ice plant)
A:Reference number: S26236; MUID:93043032; PMID:1421145
A:Accession: S26236
A:Molecule type: DNA

A:Residues: 1-25 <CU2>
A:Cross-references: EMBL:X63774; MID:G19533; PIDN:CAM45309.1; PID:G19534
C:Comment: This enzyme catalyzes the carboxylation (by carbon dioxide) of phosphoenolpyruvate to oxaloacetate.
C:Genetics:

A:Gene: ppcl
A:Introns: 56/3; 188/2; 216/3; 291/1; 325/3; 354/3; 406/3; 739/3; 868/3
C:Superfamily: phosphoenolpyruvate carboxylase
C:Keywords: carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; phosphoenolpyruvate carboxylase

Query Match 40.7%; Score 48; DB 1; Length 966;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 6; Indels

Qy 2 EWRNKKRSDWLSMVLR 17
: : : : :
Dp 475 DWTEEKRODWLLSEL 490

RESULT 5

F97772
hypothetical protein pheT [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C:Accession: F97772
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; S
cience 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowaz
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: F97772
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-818 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAU03120.1; PID:gl5619664; GSPDB:GN00
C:Genetics:
A:Gene: pheT
C:Superfamily: phenylalanine-tRNA ligase beta chain

Query Match 39.8%; Score 47; DB 2; Length 818;
Best Local Similarity 42.1%; Pred. No. 49;
Matches 8: Conservative 4: Mismatches 7: Indels

Qy 4 RNKCRSDWLSMVLRTAGVE 22
: || ||| :| : ||:
Db 225 KNKPSPDWLRKLLKNVGVK 243

RESULT 6

F69373
L-carnitine dehydratase (caIB-2) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69373
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.P.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dolis
.: Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.
Glocke, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlbeck, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A59250; MUID:98049343; PMID:9389475
A:Accession: F69373
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Molecule type: DNA
A:Residues: 1-323 <KLE>
A:Cross-references: GB:AF001036; GB:A5000782; NID:G26893359; PIDN:AAB50253.1; PID:G26491

Query Match	39.4%	Score 46.5;	DB 2;	Length 323;
Best Local Similarity	45.0%	Pred. No. 23;		
Matches	9;	Conservative	6;	Mismatches 4;
				Indels 1;
				Gaps 1;

QY 2 EWRNKKRSDWLSMVLRTAGV 21
: : | | | : : : | | |
Db 226 KFEKPRDDWIEILLK-AGV 24

RESIT.T 7

hypothetical protein RV3252c - Mycobacterium tuberculosis (strain H37RV)
F70593
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70593

R.; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70593

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-416 <COL>

A:Cross-references: GB:295121; GB:AL123456; NID:g3261742; PIDN:CAB08323.1; PID:g2072719

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV3252c

C:Superfamily: alkane 1-monooxygenase

Query Match 39.4%; Score 46.5; DB 2; Length 416;
Best Local Similarity 52.6%; Pred. No. 29;
Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 2 EWRNKKRSDWL-SMVLRTA 19
||||| : : : :
Db 20 EWRDKRYLWMLGLIAPTA 38

RESULT 8

AB0446
probable exported protein YPO3664 [imported] - *Yersinia pestis* (strain CO92)

C:Species: *Yersinia pestis*

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001

C:Accession: AB0446

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; Genotarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-638 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC93134.1; PID:gl5981586; GSPDB:GN00175

C:Genetics:

A:Gene: YPO3664

C:Superfamily: *Escherichia coli* hypothetical 73.3K protein (mreb-accB intergenic region)

Query Match 39.4%; Score 46.5; DB 2; Length 638;
Best Local Similarity 52.9%; Pred. No. 45;
Matches 9; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 5 NKXKRSWLSMVLRTAGV 21
||||| : : : :
Db 59 NKXK-EWLPWMQTLGV 74

RESULT 9

F87102
probable phosphoesterase (EC 3.1.1.-) MLI548 [similarity] - *Mycobacterium leprae*

C:Species: *Mycobacterium leprae*

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 02-Nov-2001

C:Accession: F87102

R: Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holroyd, R.; Davies, R.N.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S. Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, S.

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86903; MUID:21128732; PMID:11234002

A:Accession: F87102

A:Molecule type: DNA

A:Residues: 1-321 <STO>

A:Cross-references: GB:AL450380; NID:gl3093367; PIDN:CAC30499.1; GSPDB:GN00147

C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.

C:Genetics:

A:Gene: MLI548

C:Superfamily: *Mycobacterium tuberculosis* probable phosphoesterase RV2795c; phosphoesterase

C:Keywords: hydrolase

F:13-78/Domain: phosphoesterase core homology <PEC>

Query Match 39.0%; Score 46; DB 1; Length 321;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 EWRNKKRSDWLSMVL 16
||||| : : : :
Db 267 EWRRRXPYSWLQVL 281

RESULT 10

B70885
probable phosphoesterase (EC 3.1.1.-) RV2795c [similarity] - *Mycobacterium tuberculosis*

C:Species: *Mycobacterium tuberculosis*

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Nov-2001

C:Accession: B70885

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70885

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-324 <COL>

A:Cross-references: GB:AL008967; GB:AL123456; NID:g3261491; PIDN:CAA15590.1; PID:ell1739

A:Experimental source: strain H37RV

C:Comment: This sequence has motifs characteristic of a variety of phosphoesterases.

C:Genetics:

A:Gene: RV2795c

C:Superfamily: *Mycobacterium tuberculosis* probable phosphoesterase RV2795c; phosphoesterase

C:Keywords: hydrolase

F:16-81/Domain: phosphoesterase core homology <PEC>

Query Match 39.0%; Score 46; DB 1; Length 324;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 EWRNKKRSDWLSMVL 16
||||| : : : :
Db 270 EWRRRXPYSWLQVL 284

RESULT 11

T30935
reverse transcriptase - Atlantic horseshoe crab retrotransposon R (fragment)

C:Species: *Limulus polyphemus* (Atlantic horseshoe crab)

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30935

R: Burke, W.D.; Malik, H.S.; Jones, J.P.; Eickbush, T.H. Mol. Biol. Evol. 16, 502-511, 1999

A:Title: The domain structure and retrotransposition mechanism of R2 elements are conserved

A:Reference number: Z20937; MUID:99261661; PMID:10331276

A:Accession: T30935

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1168 <BUR>

A:Cross-references: EMBL:AF015814; NID:g3559771; PID:g3559772; PIDN:AAC34904.1

C:Genetics:

A:Mobile element: retrotransposon R

Query Match 39.0%; Score 46; DB 2; Length 1168;
Best Local Similarity 40.0%; Pred. No. 99;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 WRNKKRSDWLSMVLRTAGVE 22
||: :| | : | |
DD 919 WRDMRRSWERLALFGQGV 918

RESULT 12

T39788
 very hypothetical protein SPBC1921.04c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39788
 R:Seeger, K.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21816
 A:Accession: T39788

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1..117 <SEE>
A;Cross-references: EMBL:AL122033; PIN: CAB58970.1; GSPDB:GN00067; SPDB:SPBC1921.04C
A;Experimental source: strain 972h-, cosmid c1921
C;Genetics:
C;Gene: SPDB:SPBC1921.04C
C;Map position: 2
C;Superfamily: Schizosaccharomyces very hypothetical protein SPBC1921.04C

Query Match	38.1%;	Score 45;	DB 2;	Length 117;
Best Local Similarity	40.9%;	Pred. NO. 13;		
Matches	9;	Conservative	4;	Mismatches
			5;	Indels
			4;	Gaps
				1;

QY 3 WPNKRSD---WLSMVLRTAG 20
| : || : | : ||
DB 52 WKKKSEOEHELWASVYNTIG 73

RESULT 13

A13391
hypothetical cytosolic protein BME1119 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
R;Accession: A13391
R;Defeucheno, V.G.; Kapstrahl, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, E.S.; Mazur, M.; Gottsman, E.; Selkov, S.; Elzser, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: A13391
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-188 <GR>
A;Cross-references: KUB:A0808917; PIDN:AA152300.1; PID:g17983091; GSFDB:GN00190
A;Experimental source: strain 16M
C:Genetics:
A;Gene: BME1119
A;Map position: 1

Query Match	38.1%;	Score 45;	DB 2;	Length 188;
Best Local Similarity	38.9%;	Pred. No. 22;		
Matches	7;	Conservative	3;	Mismatches
			8;	Indels
			0;	Gaps

Qy 2 EWRNKKRSDWLSMVLRTA 19
||| ||| : :
Db 40 EWSKPVREDWIGEVKAA 57

RESULT 14

hypothetical protein 3 - phage phi-C31
C:Species: Phage phi-C31
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
C:Accession: S22568
C:Rausch, H.; Lehmann, M.
Nucleic Acids Res. 19, 5187-5189, 1991
A:Title: Structural analysis of the actinophages phiC31 attachment site.

NUCLEIC ACIDS RES. 19, 3187-3189, 1991
A;Title: Structural analysis of the actinophage phiC31 attachment site.

A,Reference number: S22556; MUID:92020203; EMID:156389
A,Accession: S22568
A,Status: preliminary; nucleic acid sequence not shown;
A,Molecule type: DNA
A,Residues: 1-132 <RAU>
A,Cross-references: EMBL:X57036
A,Note: the nucleotide sequence was submitted to the EMBL

Query Match 37.7%; Score 44.5; DB 2; Length 132;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

```

QY      1 MEWRNKRS-DWLSMVLRT 18
      :|||:::||:|:|
DB     31 LEWRSRAAYDWLAHCLOT 49

```

RESULT 15

S16055
 hypothetical protein 3 - phage phi-C31
 C:Species: phage phi-C31
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
 C:Accession: S16055
 R:Rausch, H.

R; RAUBEN, H.
submitted to the EMML Data Library December 1990

Submitted to the EMBL Data Library, December 1990

A:Reference number: S16053
A:Accession: I6055
A:Title: Problem
A:Model problem
A:Material: DNA
A:Residues: 1112
A:Cross-references: EMBL:X57036, NID:g14810, PIDN:CAA40353.1, PID:G579070
C:Genetics
A:Start codon: GTG

Query Match	37.7%	Score 44.5;	DB 2;	Length 132;
Best Local Similarity	42.1%	Pred. No. 18;		
Matches	8;	Conservative	6;	Mismatches 4;
				Indels 1;
				Gaps 1;

QY 1 MEWRNKKRS-DWLSMVLRT 18
: : : : :
Db 31 LEWRSRAAYDWLAHCLOT 49

Search completed: September 9, 2004, 17:55:14
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK: protein - protein search, using sw model

Run on: September 9, 2004, 17:39:26 ; Search time 24 Seconds
(without alignments)
47.731 Million cell updates/sec

Title: US-09-509-482D-21

Perfect score: 118

Sequence: 1 MEWRNKRSDWLSVLTAGVE 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	43.2	960	CAP1_SORBI	P29195 sorghum bic
2	48	40.7	966	CAP1_MESCR	P10490 mesembryant
3	47	39.8	818	SYFB_RICCN	Q92138 rickettsia
4	46	39.0	185	YLF1	Q87D84 xylella fas
5	45	38.1	299	ICE1_SPOPR	P89116 spodoptera
6	44.5	37.7	326	TGL2_YEAST	P54857 saccharomyc
7	44	37.3	555	DNLI_ARCFU	Q29632 archaeoglob
8	44	37.3	715	HEL5_PYPAB	Q9V049 pyrococcus
9	44	37.3	970	CAP1_MAIZE	P04711 zea mays (m
10	43	36.4	207	BRX1_CHICK	Q9W6D8 gallus gall
11	43	36.4	225	BRX1_HUMAN	Q9HBU1 homo sapien
12	43	36.4	247	BRXB_CHICK	Q9DED6 gallus gall
13	43	36.4	254	BRX1_MOUSE	Q9ER42 mus musculu
14	43	36.4	736	XMS1_DROME	Q9A3V8 drosophila
15	43	36.4	815	SYFB_RICPR	Q9ZDB4 rickettsia
16	43	36.4	880	YLB6_YEAST	Q06708 saccharomyc
17	43	36.4	960	CAP2_MESCR	P16097 mesembryant
18	43	36.4	966	CAP2_FLATR	P30694 flaveria tr
19	43	36.4	966	CAPP_FLAAU	Q42730 flaveria au
20	43	36.4	966	CAPP_MEDSA	Q02735 medicago sa
21	43	36.4	967	CAP1_FLAPR	Q01647 flaveria pr
22	43	36.4	967	CAP1_FLATR	Q01648 flaveria tr
23	43	36.4	967	CAPP_SOYBN	P51061 glycine max
24	43	36.4	967	CAPP_PEA	P51062 pisum sativ
25	43	36.4	968	CAPP_PHAVU	Q9HU12 phaseolus v
26	42.5	36.0	267	IP2A_HALN1	Q9HT88 halobacteri
27	42.5	36.0	527	TP6B_PYPAB	Q8ZVM0 pyrobaculum
28	42.5	36.0	842	PHSH_VICFA	P53537 vicia faba
29	42	35.6	205	LFTR_RHILLO	Q98NC4 rhizobium l
30	42	35.6	255	YF37_MYCTU	Q10519 mycobacteri
31	42	35.6	320	MRWV_YERPE	Q8ZIF7 yersinia pe
32	42	35.6	342	YQ6G_CAREEL	P02975 caenorhabdi
33	42	35.6	434	INGK_ECOLI	P22937 escherichia

34	42	35.6	785	1	YG51_YEAST	P50089 saccharomyc
35	42	35.6	794	1	HMAR_MOUSE	Q00547 mus musculu
36	42	35.6	918	1	EXL3_MOUSE	Q9WRL6 mus musculu
37	42	35.6	919	1	EXL3_HUMAN	O43909 h exocostin
38	42	35.6	964	1	CAPP_AWHP	Q43299 amaranthus
39	41	34.7	353	1	YN47_YERPE	Q8ZE38 yersinia pe
40	41	34.7	354	1	YCJF_YEREN	O9F4G9 yersinia en
41	41	34.7	497	1	GLPK_PUSNN	Q8RH29 fusobacteri
42	41	34.7	715	1	HEL5_PYPAB	O59025 pyrococcus
43	41	34.7	967	1	CAP1_SOYBN	O02909 glycine max
44	41	34.7	1345	1	YH00_YEAST	P38800 saccharomyc
45	40.5	34.3	135	1	ATPE_EUGGR	P31477 euglena gra

ALIGNMENTS

RESULT 1

CAP1_SORBI
ID CAP1_SORBI STANDARD; PRT; 960 AA.
AC P29195;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxylase 1 (EC 4.1.1.31) (PEPCase 1) (CP21).
GN PEPC.
OS Sorghum bicolor (Sorghum) (Sorghum vulgare).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
OX NCBI_TaxID=4558;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RX MEDLINE=91216449; PubMed=2022326;
RA Cretin C., Santi S., Keryer E., Lepiniec L., Tagu D., Vidal J.,
RA Gadal P.;
RT "The phosphoenolpyruvate carboxylase gene family of Sorghum: promoter
structures, amino acid sequences and expression of genes.";
RL Gene 99187-94(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93184205; PubMed=8443342;
RA Lepiniec L., Keryer E., Philippe H., Gadal P., Cretin C.;
RT "Sorghum phosphoenolpyruvate carboxylase gene family: structure,
function and molecular evolution.";
RL Plant Mol. Biol. 21:487-502(1993).
CC -!- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
it forms oxaloacetate, a four-carbon dicarboxylic acid source for
the tricarboxylic acid cycle.
CC -!- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
phosphoenolpyruvate + CO(2).
CC -!- ENZYME REGULATION: By light-reversible phosphorylation (By
similarity).
CC -!- PATHWAY: Tricarboxylic acid cycle. This isozyme is involved in C3
photosynthesis.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the PEPCase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X55664; CAA39197.1; -;
CC EMBL: X65137; CAA46287.1; -;
CC PIR: JH0381; JH0381.
CC PIR: S31153; S31153.
CC HSP: P00864; 1F1Y.

DR InterPro: IPR001449; PEPcase.
 DR Pfam: PF00311; PEPcase; 1.
 DR PRINTS: PR00150; PEPcase; 1.
 DR PROSITE: PS00393; PEPcase; 2; 1.
 DR PROSITE: PS00781; PEPcase; 1; 1.
 KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
 KW Tricarboxylic acid cycle; Photosynthesis; Phosphorylation.
 FT MOD_RES 7 7 PHOSPHORYLATION (BY SIMILARITY).
 FT ACT_SITE 168 168 BY SIMILARITY.
 FT ACT_SITE 596 596 BY SIMILARITY.
 SQ SEQUENCE 960 AA; 109438 MW; 73927283CE3298AA CRC64;
 Query Match 43.2%; Score 51; DB 1; Length 960;
 Best Local Similarity 56.2%; Pred. No. 4.9;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 2 EWRNKKRSDWLSVLR 17
 DB 470 ENSEKQDWLLSEL 485
 RESULT 2
 CAP1_MESCR STANDARD; PRT; 966 AA.
 ID F10490;
 AC 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DE Phosphoenolpyruvate carboxylase 1 (EC 4.1.1.31) (PEPcase 1).
 GN PPCA OR PPCI
 OS Mesembryanthemum crystallinum (Common ice plant).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Alzooceae; Mesembryanthemum.
 ON NCBI_TaxID=3544;
 RX MEDLINE=89218954; PubMed=2710107;
 RA Rickers J., Cushman J., Michalowski C., Schmitt J., Bohnert H.J.;
 RT "Expression of the CAM-form of phosphoenolpyruvate carboxylase and
 RT nucleotide sequence of a full length cDNA from Mesembryanthemum
 RT crystallinum.";
 RL Mol. Gen. Genet. 215:447-454(1989).
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89386017; PubMed=2780306;
 RA Cushman J.C., Bohnert H.J.;
 RT "Nucleotide sequence of the gene encoding a CAM specific isoform of
 RT phosphoenolpyruvate carboxylase from Mesembryanthemum crystallinum.";
 RL Nucleic Acids Res. 17:6745-6745(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE: Leaf, and Stem;
 RX MEDLINE=92393408; PubMed=2535520;
 RA Cushman J.C., Meyer G., Michalowski C.B., Schmitt J.M.,
 RA Bohnert H.J.;
 RT "Salt stress leads to differential expression of two isoenzymes of
 RT phosphoenolpyruvate carboxylase during Crassulacean acid metabolism
 RT induction in the common ice plant.";
 RL Plant Cell 3:715-725(1989).
 CC -!- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
 CC it forms oxaloacetate, a four-carbon dicarboxylic acid source for
 CC the tricarboxylic acid cycle.
 CC -!- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
 CC phosphoenolpyruvate + CO(2).
 CC -!- PATHWAY: Tricarboxylic acid cycle.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- INDUCTION: By salt stress.
 CC -!- MISCELLANEOUS: This isozyme is from salt-induced (CAM) plants.
 CC -!- SIMILARITY: Belongs to the PEPcase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X13660; CAA31356.1; -
 DR EMBL; X14587; CAA32727.1; -
 DR PIR; S05506; QYX1.
 DR HSP; P00864; 1FIY.
 DR InterPro: IPR001449; PEPcase.
 DR Pfam: PF00311; PEPcase; 1.
 DR PRINTS: PR00150; PEPcase; 1.
 DR PROSITE: PS00393; PEPcase; 2; 1.
 DR PROSITE: PS00781; PEPcase; 1; 1.
 KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
 KW Tricarboxylic acid cycle; Phosphorylation; Photosynthesis.
 FT MOD_RES 11 11 PHOSPHORYLATION (BY SIMILARITY).
 FT ACT_SITE 173 173 BY SIMILARITY.
 FT ACT_SITE 501 501 BY SIMILARITY.
 SQ SEQUENCE 966 AA; 110659 MW; 4B49C9DDB2PAD360 CRC64;
 Query Match 40.7%; Score 48; DB 1; Length 966;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 2 EWRNKKRSDWLSVLR 17
 DB 475 DWTEKQDWLLSEL 490
 RESULT 3
 SYFB_RICCN STANDARD; PRT; 818 AA.
 ID Q92138;
 AC 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Phenylalanine--tRNA synthetase beta chain (EC 6.1.1.20)
 DE (Phenylalanine--tRNA ligase beta chain) (PHERS).
 GN PHER OR RC0582.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 ON NCBI_TaxID=781;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA (Phe) = AMP +
 CC diphosphate + L-phenylalanyl-tRNA (Phe).
 CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
 CC family. Subfamily 1.
 CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL; AE008619; AAL03120.1; -
DR PIR; F97772; F97772.
DR HAMAP; MF_00283; -; 1.
DR InterPro; IPR005146; B3_4.
DR InterPro; IPR005147; B5-
DR InterPro; IPR005121; Fdx-AnticB.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004532; Phet_Bact.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF03483; B3_4; 1.
DR Pfam; PF03484; B5; 1.
DR Pfam; PF03147; PDX-ACB; 1.
DR Pfam; PF01588; tRNA_bind; 1.
DR TIGRfams; TIGR00472; phet_bact; 2.
DR PROSITE; PS00886; TRBD; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; RNA-binding; tRNA-binding;
KW Complete proteome.
FT DOMAIN 39 148 tRNA-BINDING.
FT METAL 476 476 MAGNESIUM (BY SIMILARITY).
FT METAL 482 482 MAGNESIUM (VIA CARBONYL OXYGEN)
FT METAL 485 485 (BY SIMILARITY).
FT METAL 486 486 MAGNESIUM (BY SIMILARITY).
FT METAL 486 486 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 818 AA; 90834 MW; 97FBA3DBA7B6082C CRC64;

Query Match 39.8%; Score 47; DB 1; Length 818;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 RNRKSRDLSMLVLTAGVE 22
D 225 KXKSPDWLRKLLKXNVGVK 243

RESULT 4
Y802_XYLF
ID Y802_XYLF STANDARD; PRT; 185 AA.
AC Q87DB4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0149 protein PD0802.
GN PD0802.
OS Xylella fastidiosa (strain Temecula) / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183150;
ZN [1];
SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos B.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carver H., Carraro D.W., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Martino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Penille R.C., Ferro J.A., Formighieri E.F., Kishi I.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Saseaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Seubal J.C.,
RA Kitajima J.P.;
RA "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RT J. Bacteriol. 185:1018-1026(2003).
RL J. Bacteriol. 185:1018-1026(2003).
CC -1- SIMILARITY: Belongs to the UPF0149 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U81510; AAC47442.1; -
CC HSP8; P42574; ICP3.
CC MEROPS; C14.015; -
CC InterPro; IPR002138; ICE_p10.
CC InterPro; IPR001309; ICE_p20.
CC InterPro; IPR002398; Peptidase_C14.
CC Pfam; PF00656; Peptidase_C14; 1.
CC PRINTS; PR00376; ILIBENZYM.
CC SMART; SM00115; CASC; 1.
CC PROSITE; PS01122; CASPASE_CYS; 1.
CC PROSITE; PS01121; CASPASE_HIS; 1.
CC PROSITE; PS0207; CASPASE_p10; 1.
CC PROSITE; PS0208; CASPASE_p20; 1.
CC PROSITE; PS0208; CASPASE_p20; 1.

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; AE012556; AAO28670.1; -
DR HAMAP; MF_00346; -; 1.
DR InterPro; IPR005356; UPF0149.
DR Pfam; PF03695; UPF0149; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 185 AA; 19656 MW; 354F2C8870F0D42D CRC64;

Query Match 39.0%; Score 46; DB 1; Length 185;
Best Local Similarity 66.7%; Pred. No. 5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 DWLSMVLRTAGV 21
D 46 DWLAMLADAGV 57

RESULT 5
ICE1_SPOFR
ID ICE1_SPOFR STANDARD; PRT; 299 AA.
AC P89116;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Caspase-1 precursor (SC 3.4.22.-).
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Noctuoidea;
OC Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7108;
ZN [1];
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=97153084; PubMed=8999805;
RA Ahmad M., Srinivasula S.M., Wang L., Litwack G., Fernandes-Alnemri T.,
RA Alnemri E.S.;
RT "Spodoptera frugiperda caspase-1, a novel insect death protease that
RT cleaves the nuclear immunophilin FKBP46, is the target of the
RT baculovirus antiapoptotic protein p35."
RL J. Biol. Chem. 272:1421-1424(1997).
CC -1- FUNCTION: Involved in the activation cascade of caspases
CC responsible for apoptosis execution (By similarity). Inhibited by
CC the baculovirus anti-apoptotic protein p35. Cleaves p35 and
CC nuclear immunophilin FKBP46.
CC -1- SUBUNIT: Heterodimer of a 19/18 kDa (p19/p18) and a 12 kDa (p12)
CC subunit.
CC -1- PTM: AN AUTOCATALYTIC MECHANISM GENERATES THE TWO ACTIVE SUBUNITS.
CC -1- SIMILARITY: Belongs to peptidase family C14.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U81510; AAC47442.1; -
CC HSP8; P42574; ICP3.
CC MEROPS; C14.015; -
CC InterPro; IPR002138; ICE_p10.
CC InterPro; IPR001309; ICE_p20.
CC InterPro; IPR002398; Peptidase_C14.
CC Pfam; PF00656; Peptidase_C14; 1.
CC PRINTS; PR00376; ILIBENZYM.
CC SMART; SM00115; CASC; 1.
CC PROSITE; PS01122; CASPASE_CYS; 1.
CC PROSITE; PS01121; CASPASE_HIS; 1.
CC PROSITE; PS0207; CASPASE_p10; 1.
CC PROSITE; PS0208; CASPASE_p20; 1.

```

```

DR SGD; S0002465; TGL2.
DR GO; GO:0004806; Fattyacylglycerol lipase activity; IDA.
DR Inter-Pro; IPR000379; Ser esters.
DR Hydrolase; Lipid degradation.
DR ACT SITE 144 144 CHARGE RELAY SYSTEM (BY SIMILARITY).
DR FT CONFLICT 300 300 R->H (IN REF.1).
DR FT CONFLICT 300 300
DR SQ SEQUENCE 326 AA; 37500 MW; 3D2421611ED72CE9 CRC64;

Query Match 37.7%; Score 44.5; DB 1; Length 326;
Best Local Similarity 42.9%; Pred. NO.16;
Matches 9; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

OY 1 MEWNNKRSDWLSMKVLTAGV 21
DB 285 INWKKLQDDW-SKFRRTTV 304

RESULT 7
DNL1_ARCFU
ID_DNL1_ARCFU STANDARD; PRT; 555 AA.
AC Q29632;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA ligase [EC 6.5.1.1] (polydeoxyribonucleotide synthase [ATP])..
DE LIG OR AF0623.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
ON NCBI_TaxID=2234;
RX [1]
SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RP MEDLINE=98049343; PubMed=9389475;
RY Klenk H.-P., Clayson R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richards D.L., Kurlavage A.R., Graham D.E., Kyriades N.C.,
RA Frieschmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Ariach P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997)
CC -!- FUNCTION: This protein seals, during DNA replication, DNA
CC recombination and DNA repair, nicks in double-stranded DNA (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + {deoxyribonucleotide}(N) +
CC {deoxyribonucleotide}(M) = AMP + diphosphate +
CC {deoxyribonucleotide}(N+M).
CC -!- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001061; RAAB0616.1; ALT_INIT.
CC HSSP; P00969; LA01.
CC TIGR; AF0623; -.
CC HAMAP; MF 00407; -.
CC InterPro; IPR000977; DNA_ligase.
CC Pfam; PF01068; DNA_ligase; 1.
CC Pfam; PF04679; DNA_ligase_A_C; 1.
CC Pfam; PF04675; DNA_ligase_A_N; 1.
CC TIGRFAMs; TIGR00574; dnll; 1.
CC

```



```

DR PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR PROSITE; PS00333; DNA_LIGASE_A2; FALSE_NEG.
DR PROSITE; PS00160; DNA_LIGASE_A3; 1.
KW DNA repair; DNA replication; DNA recombination; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT BINDING 249 249 AMP (BY SIMILARITY).
SQ SEQUENCE 555 AA; 63552 MW; D050FB082341EDB CRC64;

Query Match 37.3%; Score 44; DB 1; Length 555;
Best Local Similarity 56.7%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EWRNKKRSDWLS 13
Db ||| ||| |||
433 EWGEGKRSWLS 444

RESULT 8
ID HELS_PVRAB STANDARD; PRT; 715 AA.
AC Q9V0A9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative ski2-type helicase (EC 3.6.1.-).
GN PYRAB08810 OR PAB0592.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Plament D., Galperin M., Hellig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissbach J., Zivanovic Y., Porterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi.";
RL Mol. Microbiol. 47:1495-1512(2003).
CC -!- SIMILARITY: Belongs to the Helicase family. SKI2 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248285; CAB49795.1; .
DR PIR; B75135; B75135.
DR HAMAP; MF_00442; -. 1.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000445; HhH.
DR InterPro; IPR003583; HHH_1.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00633; HHH; 2.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00278; HHH1; 2.
KW Hypothetical protein; Hydrolase; Helicase; ATP-Binding;
KW Complete proteome.
FT NP BIND 46 53 ATP (POTENTIAL).
FT SITE 145 148 DEXH BOX.
SQ SEQUENCE 715 AA; 82005 MW; 611401B623690EF4 CRC64;

Query Match 37.3%; Score 44; DB 1; Length 715;
Best Local Similarity 55.6%; Pred. No. 44;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

```

```

Qy 2 EWRNKK--RSDWLSMVL 17
Db ||| ||| ||| |||
187 EWLNAKLIRSDWRPVKLR 204

RESULT 9
ID CAP1_MAIZE STANDARD; PRT; 970 AA.
AC P04711;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoenolpyruvate carboxylase 1 (EC 4.1.1.31) (PEPCase 1).
GN PEP1 OR PPC.
OS Zea mays (Maize).
OC Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73; TISSUE=Leaf;
RA Huddspeth R.L., Grula J.W.;
RT "Structure and expression of the maize gene encoding the
RT phosphoenolpyruvate carboxylase isozyme involved in C4
RT photosynthesis.";
RL Plant Mol. Biol. 12:579-589(1989).
RN [2]
RP SEQUENCE OF 39-970 FROM N.A.
RX MEDLINE=86148496; PubMed=3005978;
RA Izui K., Ishijima S., Yamaguchi Y., Katagiri F., Murata T.,
RA Shigesada K., Sugiyama T., Katsuki H.;
RT "Cloning and sequence analysis of cDNA encoding active
RT phosphoenolpyruvate carboxylase of the C4-pathway from maize.";
RL Nucleic Acids Res. 14:1615-1628(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Golden cross Bantam;
RX MEDLINE=89276342; PubMed=2731539;
RA Matsuoka M., Minami E.;
RT "Complete structure of the gene for phosphoenolpyruvate carboxylase
RT from maize.";
RL Eur. J. Biochem. 181:593-598(1989).
RN [4]
RP SEQUENCE OF 1-3 FROM N.A.
RC STRAIN=cv. H84; TISSUE=Leaf;
RX MEDLINE=90186704; PubMed=2628434;
RA Yanagisawa S., Izui K.;
RT "Maize phosphoenolpyruvate carboxylase involved in C4 photosynthesis:
RT nucleotide sequence analysis of the 5' flanking region of the gene.";
RL J. Biochem. 106:982-987(1989).
RN [5]
RP SEQUENCE OF 1-82 FROM N.A.
RX MEDLINE=8152202; PubMed=2894322;
RA Yanagisawa S., Izui K., Yamaguchi Y., Shigesada K., Katsuki H.;
RT "Further analysis of cDNA clones for maize phosphoenolpyruvate
RT carboxylase involved in C4 photosynthesis. Nucleotide sequence of
RT entire open reading frame and evidence for polyadenylation of mRNA at
RT multiple sites in vivo.";
RL FEBS Lett. 229:107-110(1988).
RN [6]
RP ACTIVE SITE, AND SEQUENCE OF 599-610.
RX MEDLINE=91098247; PubMed=2268676;
RA Jiao J.-A., Podesta F.E., Chollet R., O'Leary M.H., Andreo C.S.;
RT "Isolation and sequence of an active-site peptide from maize leaf
RT phosphoenolpyruvate carboxylase inactivated by pyridoxal
RT 5'-phosphate.";
RL Biochim. Biophys. Acta 1041:291-295(1990).
RN [7]
RP PHOSPHORYLATION SITE.
RA Jiao J.-A., Vidal J., Echevarria C., Chollet R.;
RT "In vivo regulatory phosphorylation site in C4-leaf
RT phosphoenolpyruvate carboxylase from maize and sorghum.";

```

```

RL Plant Physiol. 96:297-301(1991).
CC -!- FUNCTION: Through the carboxylation of phosphoenolpyruvate (PEP)
CC it forms oxaloacetate, a four-carbon dicarboxylic acid source for
CC the tricarboxylic acid cycle.
CC -!- CATALYTIC ACTIVITY: Phosphate + oxaloacetate = H(2)O +
CC phosphoenolpyruvate + CO(2).
CC -!- ENZYME REGULATION: By light-reversible phosphorylation.
CC -!- PATHWAY: Tricarboxylic acid cycle. This isozyme is involved in C4
CC photosynthesis.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the PEPCase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15238; CAA33316.1; -
DR EMBL; X03613; CAA27270.1; -
DR EMBL; X14581; CAA32724.1; -
DR EMBL; X14579; CAA32724.1; ALT INIT.
DR EMBL; X14580; CAA32723.1; -
DR EMBL; X15642; CAA33663.1; -
DR EMBL; X07168; CAA30158.1; -
DR PDB; 1UOQ; 14-JAN-03.
DR MaizeDB; 30066; -.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PR00150; PEPCaseXKLASE.
DR PROSITE; PS00393; PEPCase_2; 1.
DR PROSITE; PS00781; PEPCase_1; 1.
KW Lyase; Carbon dioxide fixation; Allosteric enzyme; Multigene family;
KW Tricarboxylic acid cycle; Phosphorylation; Photosynthesis;
KW 3D-structure.
FT MOD_RES 15 15 PHOSPHORYLATION.
FT ACT_SITE 177 177 BY SIMILARITY.
FT ACT_SITE 606 606 A -> D (IN REF. 2 AND 3).
FT CONFLICT 239 239 EL -> DV (IN REF. 2).
FT CONFLICT 338 339 P -> S (IN REF. 2 AND 3).
FT CONFLICT 482 482 D -> E (IN REF. 3).
FT CONFLICT 509 509 QPL -> PAV (IN REF. 2 AND 3).
FT CONFLICT 557 559 D -> S (IN REF. 2 AND 3).
FT CONFLICT 570 574 SA -> LR (IN REF. 2).
FT CONFLICT 573 574 C -> S (IN REF. 2).
FT CONFLICT 687 687 A -> P (IN REF. 2).
FT CONFLICT 726 726 A -> R (IN REF. 2).
FT CONFLICT 963 963 A -> R (IN REF. 2).
SQ SEQUENCE 970 AA; 109296 MW; 95B66F96ABCE22F4 CRC64;
Query Match 37.3%; Score 44; DB 1; Length 970;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 2 EWRNKRSDWLSMVL 17
Db 480 EWPEDKQEWELSEL 495

```

```

RESULT 10
BRX1 CHICK
ID BRX1 CHICK STANDARD; PRT; 207 AA.
AC Q9W5D8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein BarH-like 1 (Fragment).
GN BARX1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC Gallus.
CC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=99228399; PubMed=10213385;
RA Barlow A.J., Bogard J.P., Ladher R., Francis-West P.H.;
RT "Expression of chick barx-1 and its differential regulation by FGF-8
RT and BMP signaling in the maxillary primordia."
RL Dev. Dyn. 214:291-302(1999).
CC -!- FUNCTION: Transcription factor, which may be involved in
CC craniofacial development, in odontogenesis and in stomach
CC organogenesis. May have a role in the differentiation of molars
CC from incisors. Binds to a regulatory module of the NCAM promoter.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: Expressed predominantly in the facial
CC primordia, developing stomach, and proximal limbs.
CC -!- DEVELOPMENTAL STAGE: First detectable in the facial primordia at
CC stage 18 after neural crest migration. Expressed in regions
CC derived from both mid- and hindbrain neural crest. Also expressed
CC in the developing cartilage elements of the limb, first within a
CC restricted population in the prechondrogenic mesenchyme and later
CC in the rounded chondrocytes at the epiphyses of developing long
CC bones.
CC -!- SIMILARITY: Belongs to the BAR homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF116460; AAD21043.1; -
DR HSP; P14653; IB72.
DR TRANSFAC; TC9394; -.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_lambrepreser.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX
DR PRINTS; PR00031; HTHREPRESSR.
DR PRODOM; ED000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW Homeobox; Transcription regulation; DNA-binding; Nuclear protein.
FT NON_TER 1 1
FT DNA_BIND 95 154 HOMEBOX.
SQ SEQUENCE 207 AA; 22467 MW; 3ED64A91D3BCE84 CRC64;
Query Match 36.4%; Score 43; DB 1; Length 207;
Best Local Similarity 35.0%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 3 WRNKRSDWLSMVLRTAGVE 22
Db 142 WYQNRKMKWKIVLQGGLE 161

```

```

RESULT 11
BRX1 HUMAN
ID BRX1 HUMAN STANDARD; PRT; 225 AA.
AC Q9HBU1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein BarH-like 1.
GN BARX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND VARIANT ALA-19.
Tissue=craniofacial;
MEDLINE=20453194; PubMed=10995576;
Gould D.B., Walter M.A.;
"Cloning, Characterization, localization, and mutational screening of
the human BARX1 gene."
Genomics 68:336-342(2000).
-!- FUNCTION: Transcription factor, which may be involved in
craniofacial development, in ontogenesis and in stomach
organogenesis. May have a role in the differentiation of molars
from incisors. Binds to a regulatory module of the NCAM promoter.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- TISSUE SPECIFICITY: Widely expressed. Expressed at higher levels
in testis and heart. Detected in craniofacial tissue and adult
iris, but not in lymphocytes, fibroblasts, choroid retina, retinal
pigment epithelium, kidney, or fetal liver.
-!- POLYMORPHISM: The polymorphism is not associated with Axenfeld-
Reiger syndrome (ARS), iridogoniodysgenesis syndrome (IGDS) or
related ocular malformations.
-!- SIMILARITY: Belongs to the BAR homeobox family.
-!- SIMILARITY: Contains 1 homeobox domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF213356; AAG23738.1; -;
HSSP; P14653; 1B72.
Genew; HGNC:955; BARX1.
MIM; 603260; -;
GO; GO:0000228; C:nuclear chromosome; NAS.
GO; GO:0003700; P:transcription factor activity; NAS.
GO; GO:0007235; P:development; NAS.
GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
InterPro; IPR001356; Homeobox.
InterPro; IPR000047; HTH_lambdarepressor.
Pfam; PF00046; homeobox; 1.
PRINTS; PR00024; HOMEBOX.
PRINTS; PR00031; HTHREPRESSR.
ProDom; PD00010; Homeobox; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEBOX_1; 1.
PROSITE; PS50071; HOMEBOX_2; 1.
Homeobox; Transcription regulation; DNA-binding; Nuclear protein;
Polymorphism.
DNA BIND 113 172 HOMEBOX.
DOMAIN 15 24 POLY-ALA.
VARIANT 19 19 T->A.
/FTID=VAR_010927.
SEQUENCE 225 AA; 24061 MW; A7907BB4666F393 CRC64;
Query Match 36.4%; Score 43; DB 1; Length 225;
Best Local Similarity 35.0%; Pred. No. 18;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 3 WRNKKRSDWLSMVLRTAGVE 22
Db 160 WYQNRBMKWKXIVLQGGGLE 179
RESULT 12
BRXB_CHICK STANDARD; PRT; 247 AA.
AC Q9DE6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

Homeobox protein BarH-like 1b (Bar class homeoprotein BarX1b).
BARX1b.
Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
MEDLINE=21259836; PubMed=11359793;
Nakamura M., Nishida W., Mori S., Hiwada K., Hayashi K., Sobue K.;
"Transcriptional activation of beta-tropomyosin mediated by serum
response factor and a novel Barx homologue, BarX1b, in smooth muscle
cells."
J. Biol. Chem. 276:18313-18320(2001).
-!- FUNCTION: Transcription factor which is involved with the serum
response factor (SRF) in the smooth muscle cell-specific
transcription of the beta-tropomyosin gene in the upper digestive
organs and their attached arteries.
-!- SUBUNIT: Interacts with serum response factor (SRF).
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- TISSUE SPECIFICITY: Expressed in smooth muscle cells of the upper
digestive organs and their attached arteries and to craniofacial
structures.
-!- SIMILARITY: Belongs to the BAR homeobox family.
-!- SIMILARITY: Contains 1 homeobox domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AB044371; BAB18919.1; -;
HSSP; P14653; 1B72.
InterPro; IPR001356; Homeobox.
InterPro; IPR000047; HTH_lambdarepressor.
Pfam; PF00046; homeobox; 1.
PRINTS; PR00024; HOMEBOX.
PRINTS; PR00031; HTHREPRESSR.
ProDom; PD00010; Homeobox; 1.
SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEBOX_1; 1.
PROSITE; PS50071; HOMEBOX_2; 1.
Homeobox; Transcription regulation; DNA-binding; Nuclear protein.
FT DNA BIND 135 194 HOMEBOX.
SEQUENCE 247 AA; 27027 MW; A72EFA192F8624F CRC64;
Query Match 36.4%; Score 43; DB 1; Length 247;
Best Local Similarity 35.0%; Pred. No. 20;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 3 WRNKKRSDWLSMVLRTAGVE 22
Db 182 WYQNRBMKWKXIVLQGGGLE 201
RESULT 13
BRX1_MOUSE STANDARD; PRT; 254 AA.
ID BRX1_MOUSE
AC Q9ER42; O09066; P70159; Q9ERV2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Homeobox protein BarH-like 1.
GN BARX1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RA Buchner G.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 30-254 FROM N.A.
 RC STRAIN=BALE/C;
 RA Meech R., Edelman D.B., Jones F.S.;
 RT "Characterization of the mouse Barx1 gene."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 108-254 FROM N.A.
 RC STRAIN=NMRI;
 RA Tissier-Sera J.P., Mucchielli M.L., Mark M., Mattei M.-G., Goridis C.,
 RA Brunet J.F.;
 RT "Barx1, a new mouse homeodomain transcription factor expressed in
 RT cranio-facial ectomesenchyme and the stomach."
 RL Mech. Dev. 51:3-15 (1995).
 CC -!- FUNCTION: Transcription factor, which may be involved in
 CC craniofacial development, in ontogenesis and in stomach
 CC organogenesis. May have a role in the differentiation of molars
 CC from incisors. Binds to a regulatory module of the NCAM promoter.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in the facial
 CC primordia, developing stomach, and proximal limbs.
 CC -!- DEVELOPMENTAL STAGE: Expressed in areas of the first and second
 CC branchial arches, before any apparent cellular or morphologic
 CC differentiation. Later in development, all expressing tissue in
 CC this region, including the mesenchyme underlying the olfactory
 CC epithelium, the primary and secondary palate, the molar tooth
 CC papillae, and the stroma of the submandibular gland, appear to be
 CC derived from ectomesenchyme of neural crest origin. By day 16.5,
 CC all areas except the developing molars are BARX1-negative. In
 CC addition, BARX1 marks the area of the future stomach in the
 CC primitive gut at embryonic day 9.5, and is present in the
 CC mesenchymal wall of the stomach until embryonic day 16.5.
 CC -!- SIMILARITY: Belongs to the BAR homeobox family.
 CC -!- SIMILARITY: Contains 1 homeobox domain.
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-30 is the initiator.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ297677; CAC10357.1; -;
 CC EMBL; AF277160; AAG18573.1; -;
 CC EMBL; Y07960; CAA69257.1; -;
 CC HSSP; P14653; 1B72.
 CC TRANSFAC; T02403; -;
 CC MGD; MGI:10124; Barx1.
 CC InterPro; IPR001356; Homeobox.
 CC InterPro; IPR000047; HTH lambrpressor.
 CC Pfam; PF00046; homeobox; 1.
 CC PRINTS; PR00024; HOMEBOX.
 CC PRINTS; PR00031; HTHREPRESSR.
 CC ProDom; PD000010; Homeobox; 1.
 CC SMART; SM00389; HOX; 1.
 CC PROSITE; PS00027; HOMEBOX_1; 1.
 CC PROSITE; PS00071; HOMEBOX_2; 1.
 CC Homeobox; Transcription regulation; DNA-binding; Nuclear protein.
 CC DNA_BIND 142 201 HOMEBOX.
 CC DOMAIN 44 53 POLY-ALA.
 CC SEQUENCE 254 AA; 27282 MW; 48586B28F4A23FCC CRC64;
 SQ

Query Match 36.4%; Score 43; DB 1; Length 254;
 Best Local Similarity 35.0%; Pred. No. 21;
 Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Oy 3 WRNKBSDWLSWLTAGVE 22
 Db 189 WYQNRKMKWIKVIGGGLE 208
 RESULT 14
 XMS1 DROME STANDARD; PRT; 736 AA.
 ID _XMS1 DROME STANDARD; PRT; 736 AA.
 AC Q9U3V8; Q9VX76;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Xmas-1 protein.
 OS XMAS-1 OR CG32561/CG8919.
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP Xu E.Y., Kaufman T.C., Wu C.;
 RT "Two overlapping genes, xmas-1 and xmas-2, are required for
 RT spermatogenesis, oogenesis and embryogenesis";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkley;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balvel R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Fowler K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.B., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuesken D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas A., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RN Science 287:2185-2195 (2000).
 RN [3]
 RN REVISIONS.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
BL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Required for spermatogenesis, oogenesis and
CC embryogenesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF216664; AAF23814.1; -;
DR EMBL; AE003504; AAF48702.2; -;
DR FlyBase; FBGN0016080; xmas-1.
DR GO; GO:0009790; P:embryonic development; NAS.
DR GO; GO:0007292; P:female gamete generation; NAS.
DR GO; GO:0007283; P:spermatogenesis; NAS.
KW Developmental protein.
FT CONFLICT 31 31 Y -> H (IN REF. 1).
FT CONFLICT 231 231 A -> P (IN REF. 1).
FT CONFLICT 721 721 K -> E (IN REF. 1).
SQ SEQUENCE 736 AA; 84460 MW; 8D0D886578907E09 CRC64;

Query Match 36.4%; Score 43; DB 1; Length 736;
Best Local Similarity 63.6%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MEWENKGRSDW 11
Db 348 MWEEQKRSPPW 358

RESULT 15
SYBF RICPR STANDARD; PRT; 815 AA.
ID -SVPS RICPR Q9ZDE4;
AC Q9ZDE4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE {Phenylalanine--tRNA ligase beta chain} (Phers).
GN PHET OR RP418
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=98039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140 (1998).
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA (Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA (Phe).
CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
CC family. Subfamily 1.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ235271; CAA14875.1; -;
DR PIR; A71700; A71700.
DR HSP; P27002; 1EYS.
DR HAMAP; MF_00283; -; 1.
DR InterPro; IPR005146; B3_4.
DR InterPro; IPR005147; B5.
DR InterPro; IPR005121; Fdx-AntiCB.
DR InterPro; IPR008934; Nucleic_acid_OB.
DR InterPro; IPR004532; Phet_bact.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF03483; B3_4; 1.
DR Pfam; PF03484; B5; 1.
DR Pfam; PF03147; FDX-ACB; 1.
DR Pfam; PF01588; tRNA_bind; 1.
DR TIGRfam; TIGR00472; phet_bact; 2.
DR PROSITE; PS00886; TRBD; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; RNA-binding; tRNA-binding;
KW Complete proteome.
FT DOMAIN 39 148 tRNA-BINDING.
FT METAL 474 474 MAGNESIUM (BY SIMILARITY).
FT METAL 480 480 MAGNESIUM (VIA CARBONYL OXYGEN)
FT METAL 483 483 (BY SIMILARITY).
FT METAL 484 484 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 815 AA; 91584 MW; F508E95333FCAB43 CRC64;

Query Match 36.4%; Score 43; DB 1; Length 815;
Best Local Similarity 31.6%; Pred. No. 73;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 4 RNKGRSDWLSMVLRTAGVE 22
Db 225 KNKPSNWLQQLKXVGIK 243

Search completed: September 9, 2004, 17:52:30
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 17:46:32 ; Search time 113 Seconds
(without alignments)
61.428 Million cell updates/sec

Title: US-09-509-482D-21
Perfect score: 118
Sequence: 1 MERNKRSWLSVLRTAGVE 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum Match 0%
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_ricet.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	43.2	959	Q9FYX8	Q9FYX8 cryza sativ
2	51	43.2	960	Q9SAZ6	Q9SAZ6 sea maye (m
3	51	43.2	961	Q84M23	Q84M23 echinochloa
4	50	42.4	326	Q8Y209	Q8Y209 ralsconia s
5	50	42.4	565	Q9YBE0	Q9YBE0 aeropyrum p
6	49	41.5	529	Q8KNJ0	Q8KNJ0 micromonosp
7	48	40.7	186	Q8DV98	Q8DV98 streptococc
8	48	40.7	329	Q7WKC4	Q7WKC4 bordetella
9	48	40.7	329	Q7VSP1	Q7VSP1 bordetella
10	48	40.7	460	Q7VSV5	Q7VSV5 prochloroco
11	48	40.7	2268	Q814N5	Q814N5 plasmodium
12	47	39.8	242	Q83NN0	Q83NN0 tropheryma
13	47	39.8	242	Q83GA6	Q83GA6 tropheryma
14	47	39.8	794	Q98MC3	Q98MC3 rhizobium l
15	47	39.8	964	Q8S228	Q8S228 setaria ita
16	46.5	39.4	323	Q29272	Q29272 archaeoglob

ALIGNMENTS

RESULT 1

Q9FYX8 PRELIMINARY; PRT; 959 AA.
ID Q9FYX8
AC Q9FYX8
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCase).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Rhartoideae; Oryzeae; Oryza.
OK NCBI_TaxID=4530;
[1]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
RA Yamamoto N., Kurita A., Masumura T., Sugimoto T., Morita S.,
RA Shiraiishi N., Oji Y., Tanaka K.;
RT "Root type of phosphoenolpyruvate carboxylase in developing rice
seeds.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TO FORM OXALACETATE, A FOUR-CARBON DICARBOXYLIC ACID
CC SOURCES FOR THE TRICARBOXYLIC ACID CYCLE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: PHOSPHATE + OXALACETATE = H(2)O +
CC -!- PATHWAY: PHOSPHOENOLPYRUATE + CO(2).
CC -!- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC EXEL; AF271995; BELONGS TO THE PEPCASE FAMILY.
DR HSSP; P00864; 1EY.
DR Gramene; Q9FYX8;
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008964; F:phosphoenolpyruvate carboxylase activity; IEA.
DR GO; GO:0015977; P:carbon utilization by fixation of carbon di...; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PR00150; PEPCARBLXASE.
DR PROSITE; PS00783; PEPCASE_1; 1.
DR PROSITE; PS00393; PEPCASE_2; 1.

O05895 mycobacteri
Q7TW33 mycobacteri
Q8ZAW7 yersinia pe
Q90XS7 oncorhynchu
Q982B0 rhizobium l
Q81BP6 plasmodium
Q92516 mycobacteri
Q33337 mycobacteri
Q7XV0 mycobacteri
Q8IN22 drosophila
Q7UJF0 synchococc
Q81J22 x mokara cv
Q44318 limulus pol
Q83XD3 acinetobact
Q81SD5 sterculia f
Q8USX2 echinosacch
Q8YK59 brucella me
Q8G175 brucella su
Q81955 spodoptera
Q8U0N3 pyrococcus
Q8RIQ3 mus musculu
Q8TNC6 methanosarc
Q8ZME6 salmonella
Q84XR7 zea mays (m
Q8JVC1 penicillium
Q9ZX82 bacterioph
Q44589 alcaligenes
Q18960 caenorhabdi

```

KW Carbon dioxide fixation; Lyase; Pyruvate; Tricarboxylic acid cycle.
SQ SEQUENCE 959 AA; 109189 MW; D9BE1B982CAE33C4 CRC64;

Query Match 43.2%; Score 51; DB 10; Length 959;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWRNKRSDWLSWVLR 17
DB 469 EWSEKRDWLLSEL 484

RESULT 2
Q9SAZ6 PRELIMINARY; PRT; 960 AA.
AC Q9SAZ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (PEPCase).
GN PFC1C.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H4; TISSUE=Root;
RA Dong L., Masuda T., Kawamura T., Hata S., Izui K.;
RT "Cloning, expression and characterization of a root-form
RT phosphoenolpyruvate carboxylase from Zea mays: Comparison with the C4-
RT form enzyme.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TO FORM OXALACETATE. A FOUR-CARBON DICARBOXYLIC ACID
CC SOURCE FOR THE TRICARBOXYLIC ACID CYCLE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: PHOSPHATE + OXALACETATE = H(2)O +
CC PHOSPHOENOLPYRUVATE + CO(2).
CC -!- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -!- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.
DR EMBL; A3012228; BAA28170.1;
DR HSP; P00864; 1FIY.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008964; F:phosphoenolpyruvate carboxylase activity; IEA.
DR GO; GO:0015977; P:carbon utilization by fixation of carbon di. . . ; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase.
DR PRINTS; PR00150; PEPCARXKLASE.
DR PROSITE; PS00781; PEPCASE_1; 1.
DR PROSITE; PS00393; PEPCASE_2; 1.
KW Carbon dioxide fixation; Lyase; Pyruvate; Tricarboxylic acid cycle.
SQ SEQUENCE 960 AA; 109428 MW; 7B7F156A506F7C2D CRC64;

Query Match 43.2%; Score 51; DB 10; Length 960;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWRNKRSDWLSWVLR 17
DB 470 EWSEKRDWLLSEL 485

RESULT 3
Q84MZ3 PRELIMINARY; PRT; 961 AA.
AC Q84MZ3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31).
GN PFC.
OS Echinochloa crus-galli (Barnyard grass).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Paniceae; Echinochloa.
OX NCBI_TaxID=90397;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Zhang G., Zhao M., Ding Z., Zhang L.;
RT "Echinochloa crus-galli phosphoenolpyruvate carboxylase mRNA.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY251492; AAP06951.1;
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008964; F:phosphoenolpyruvate carboxylase activity; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PR00150; PEPCARXKLASE.
DR PROSITE; PS00781; PEPCASE_1; 1.
DR PROSITE; PS00393; PEPCASE_2; 1.
KW Lyase; Pyruvate
SQ SEQUENCE 961 AA; 109324 MW; 8CD24641B2D663F3 CRC64;

Query Match 43.2%; Score 51; DB 10; Length 961;
Best Local Similarity 56.2%; Pred. No. 42;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWRNKRSDWLSWVLR 17
DB 471 EWSEKRDWLLSEL 486

RESULT 4
Q8Y2U9 PRELIMINARY; PRT; 326 AA.
AC Q8Y2U9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative integral membrane transmembrane protein.
GN RSC0233 OR R300665.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RC MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646058; CAD13761.1;
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 1.
KW Complete proteome.
SQ SEQUENCE 326 AA; 34149 MW; 4D42B69813919F78 CRC64;

Query Match 42.4%; Score 50; DB 16; Length 326;
Best Local Similarity 57.1%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 WRNKRSDWLSWVLR 16
DB 79 WRKTRRDWLLALVL 92

RESULT 5

```

```

Q9YBB0
ID Q9YBB0 PRELIMINARY; PRT; 565 AA.
AC Q9YBB0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 565AA long hypothetical methymalonyl-CoA mutase alpha-subunit.
GN APE1687.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NK1.
RA MEDLINE=99310339; PubMed=10389966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80684.1; -
DR PIR; C72550; C72550.
DR HSSP; P11653; 5REQ.
DR GO; GO:000434; F:methylmalonyl-CoA mutase activity; IEA.
DR GO; GO:0004815; P:metabolism; IEA.
DR InterPro; IPR006099; MMCOA_mutase.
DR InterPro; IPR006098; MMCOA_mutase_N.
DR Pfam; PF01642; MM_COA_mutase; 1.
DR TIGRFAMs; TIGR00641; acid_COA_mut_N; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 565 AA; 64884 MW; 405BD7F5540CD6A9 CRC64;

Query Match 42.4%; Score 50; DB 17; Length 565;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 12; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 EW--RNKGRSDWLSWVLRAGV 21
Db ||||| :|||
316 EWFAGKXKRSWLRFTHTAGV 337

RESULT 6
ID Q8KNJ0 PRELIMINARY; PRT; 529 AA.
AC Q8KNJ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CALT1.
GN CALT1.
OS Micromonospora echinospora (Micromonospora purpurea).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micromonosporineae; Micromonosporaceae; Micromonospora.
OX NCBI_TaxID=1877;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 15839.
RA Allert J., Shepard E., Lonovskaya N., Zazopoulos E., Staffa A.,
RA Bachmann B.O., Huang K., Fontstein L., Czişny A., Whitwam R.E.,
RA Farnet C.M., Thorson J.S.;
RT "The calicheamicin gene cluster and its iterative type I PKS.";
RL Science 0:0-0(2002).
DR EMBL; AF497482; AAM94765.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:001520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

Q9YBB0
ID Q9YBB0 PRELIMINARY; PRT; 329 AA.
AC Q9YBB0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 329AA long hypothetical methymalonyl-CoA mutase alpha-subunit.
GN APE1687.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NK1.
RA MEDLINE=99310339; PubMed=10389966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80684.1; -
DR PIR; C72550; C72550.
DR HSSP; P11653; 5REQ.
DR GO; GO:000434; F:methylmalonyl-CoA mutase activity; IEA.
DR GO; GO:0004815; P:metabolism; IEA.
DR InterPro; IPR006099; MMCOA_mutase.
DR InterPro; IPR006098; MMCOA_mutase_N.
DR Pfam; PF01642; MM_COA_mutase; 1.
DR TIGRFAMs; TIGR00641; acid_COA_mut_N; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 565 AA; 64884 MW; 405BD7F5540CD6A9 CRC64;

Query Match 41.5%; Score 49; DB 2; Length 529;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 RNKGRSDWLSWVLRAGV 21
Db ||||| :|||
207 RKVYRIDWLGALLITSGV 224

RESULT 7
ID Q8DV98 PRELIMINARY; PRT; 186 AA.
AC Q8DV98;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN SMU 600C.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Gia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014904; AAN58339.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 186 AA; 21045 MW; 049F3724668024B9 CRC64;

Query Match 40.7%; Score 48; DB 16; Length 186;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 MEWRNKRSDWLSWVLRTA 19
Db :||| :|||
105 LSWTRKK--DWLHVTIATA 121

RESULT 8
ID Q7WKC4 PRELIMINARY; PRT; 329 AA.
AC Q7WKC4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 329AA long hypothetical methymalonyl-CoA mutase alpha-subunit.
GN APE1687.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NK1.
RA MEDLINE=99310339; PubMed=10389966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80684.1; -
DR PIR; C72550; C72550.
DR HSSP; P11653; 5REQ.
DR GO; GO:000434; F:methylmalonyl-CoA mutase activity; IEA.
DR GO; GO:0004815; P:metabolism; IEA.
DR InterPro; IPR006099; MMCOA_mutase.
DR InterPro; IPR006098; MMCOA_mutase_N.
DR Pfam; PF01642; MM_COA_mutase; 1.
DR TIGRFAMs; TIGR00641; acid_COA_mut_N; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 565 AA; 64884 MW; 405BD7F5540CD6A9 CRC64;

Query Match 42.4%; Score 50; DB 17; Length 565;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 12; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 EW--RNKGRSDWLSWVLRAGV 21
Db ||||| :|||
316 EWFAGKXKRSWLRFTHTAGV 337

RESULT 6
ID Q8KNJ0 PRELIMINARY; PRT; 529 AA.
AC Q8KNJ0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CALT1.
GN CALT1.
OS Micromonospora echinospora (Micromonospora purpurea).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micromonosporineae; Micromonosporaceae; Micromonospora.
OX NCBI_TaxID=1877;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 15839.
RA Allert J., Shepard E., Lonovskaya N., Zazopoulos E., Staffa A.,
RA Bachmann B.O., Huang K., Fontstein L., Czişny A., Whitwam R.E.,
RA Farnet C.M., Thorson J.S.;
RT "The calicheamicin gene cluster and its iterative type I PKS.";
RL Science 0:0-0(2002).
DR EMBL; AF497482; AAM94765.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR GO; GO:001520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

```



```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Integrase.
GN B32185.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RE50 / ATCC BAA-588;
RX PARKILL J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640443; CAE32681.1; -.
KW Complete proteome.
SQ SEQUENCE 329 AA; 37458 MW; 8F5AC03F4B3E4682 CRC64;

Query Match 40.7%; Score 48; DB 16; Length 329;
Best Local Similarity 42.1%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 EWRNKKRSDWLSMVLRTAG 20
DB 137 EWRNLKDDPWLDTMPPEG 155

RESULT 9
QTVSD1 PRELIMINARY; PRT; 329 AA.
AC QTVSD1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Integrase.
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN BPO502.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640412; CAE44831.1; -.
KW Complete proteome.
SQ SEQUENCE 329 AA; 37471 MW; 9FC0F1732CB7C74A CRC64;

Query Match 40.7%; Score 48; DB 16; Length 329;
Best Local Similarity 42.1%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```

```

Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 EWRNKKRSDWLSMVLRTAG 20
DB 137 EWRNLKDDPWLDTMPPEG 155

RESULT 10
QTVSV5 PRELIMINARY; PRT; 460 AA.
AC QTVSV5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UDP-N-acetylmuramoylalanine-D-glutamate ligase (EC 6.3.2.9).
GN MURD OR PWT1434.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21609.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 460 AA; 49543 MW; 88073C9A6F8DC803 CRC64;

Query Match 40.7%; Score 48; DB 16; Length 460;
Best Local Similarity 41.7%; Pred. No. 56;
Matches 10; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 1 MEWRF--NKXGSDWLSMVLRTAGV 22
DB 154 LKWRQPNQAPDPWLVLMBLSSVQIE 177

RESULT 11
Q8I4N5 PRELIMINARY; PRT; 2268 AA.
AC Q8I4N5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte membrane protein 1 (PfEMP1).
GN PF12665C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22285705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Paine A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McPadden G.I., Cummings L.M., Subramanian G.K., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AE014852; AAN36615.1; -.

```

DR GO: GO:0005539; F:glycosaminoglycan binding; IEA.
 DR GO: GO:0009405; P:patogenesis; IEA.
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP; 2.

SQ SEQUENCE 2268 AA; 255414 MW; 9B61A50B525DEC54 CRC64;

Query Match 40.7%; Score 48; DB 5; Length 2268;
 Best Local Similarity 70.0%; Pred. No. 3e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EWRNKRSDW 11
 ||| ||| ||| |||
 Db 1575 EWINKRTEW 1584

RESULT 12

Q83NMO PRELIMINARY; PRT; 242 AA.
 AC Q83NMO;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative ABC transporter integral membrane subunit.
 GN TWJ62.
 OS Tropheryma whipplei (strain TW08/27) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=218496;
 EN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22495039; PubMed=12606174;
 RA Bentley S.D., Maitwald M., Murphy L.D., Pallen M.J., Yeats C.A.,
 RA Dover L.G., Norbertczak H.T., Besta G.S., Quail M.A., Harris D.E.,
 RA von Herbay A., Goble A., Rutter S., Squares R., Squares S.,
 RA Barrell B.G., Parkhill J., Rellman D.A.;
 RT "Sequencing and analysis of the genome of the Whipple's disease
 bacterium Tropheryma whipplei.";
 RL Lancet 361:637-644(2003).
 DR EMBL; BX251411; CAD67033.1; -.
 KW Complete proteome.
 SQ SEQUENCE 242 AA; 26748 MW; DOE2BD8E531BF250 CRC64;

Query Match 39.8%; Score 47; DB 16; Length 242;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 7 KRSDWLSVLRTAGVE 22
 ::|||::|:
 Db 191 RRAYWLSIVLKRGIE 206

RESULT 13

Q83GA6 PRELIMINARY; PRT; 242 AA.
 AC Q83GA6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN TW7407.
 OS Tropheryma whipplei (strain Twist) (Whipple's bacillus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Cellulomonadaceae; Tropheryma.
 OX NCBI_TaxID=203267;
 EN [1]
 RP SEQUENCE FROM N.A.
 RA Raoult D., Audic S., Robert C., Ogata H., Suhre K., Drancourt M.,
 RA Claverie J.-M.;
 RT "Tropheryma whipplei illustrates the diversity of gene loss patterns
 in small genome bacterial pathogens";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB016851; AA04504.1; -.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 242 AA; 26748 MW; DOE2BD8E531BF250 CRC64;

Query Match 39.8%; Score 47; DB 16; Length 242;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 7 KRSDWLSVLRTAGVE 22
 ::|||::|:
 Db 191 RRAYWLSIVLKRGIE 206

RESULT 14

Q88MC3 PRELIMINARY; PRT; 794 AA.
 AC Q88MC3;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE Outer membrane protein.
 GN MLL0636.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 EN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP002995; BAB48190.1; -.
 DR InterPro; IPR000184; Bac_surfAg_D15.
 DR Pfam; PF01103; Bac_surface_Ag; 1.
 KW Complete proteome.
 SQ SEQUENCE 794 AA; 85741 MW; BABE4A1BC8FEEDC6 CRC64;

Query Match 39.8%; Score 47; DB 16; Length 794;
 Best Local Similarity 61.5%; Pred. No. 1.4e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 NKGRSDWLSVLR 17
 ||| ||| ||| |||
 Db 210 NTKRSSWVSFLR 222

RESULT 15

Q8S2Z8 PRELIMINARY; PRT; 964 AA.
 AC Q8S2Z8;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 24, Last annotation update)
 DE Putative C4 phosphoenolpyruvate carboxylase (EC 4.1.1.31)
 GN PPC.
 OS Setaria italica (Foxtail millet).
 OC Sukariyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Paniceae; Setaria.
 OX NCBI_TaxID=4555;
 EN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Gufeng; TISSUE=leaf;
 RA Ding Z., Jing Y., Lin R., Li L., Xu Y., Kuang T.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: TO FORM OXALACETATE, A FOUR-CARBON DICARBOXYLIC ACID
 CC -!- SOURCE FOR THE TRICARBOXYLIC ACID CYCLE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: PHOSPHATE + OXALOACETATE = H(2)O +
CC PHOSPHENOLPYRUVATE + CO(2).
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.
DR EMBL; AF495586; AAM15963.1; -;
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008964; P:phosphoenolpyruvate carboxylase activity; IEA.
DR GO; GO:0015977; P:carbon utilization by fixation of carbon di. . . ; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR001449; PEPCase.
DR Pfam; PF00311; PEPCase; 1.
DR PRINTS; PR00150; PEPCARBXLA.
DR PROSITE; PS00781; PEPCASE 1; 1.
DR PROSITE; PS00393; PEPCASE 2; 1.
KW Carbon dioxide fixation: Lyase; Pyruvate; Tricarboxylic acid cycle.
SQ SEQUENCE 964 AA; 110069 MW; A86E4F7330B7D10C CRC64;

Query Match 39.8%; Score 47; DB 10; Length 964;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWRNKKRSDWLSMVLK 17
||| :||| :|||
Db 474 EWSEKQEWLSELR 489

Search completed: September 9, 2004, 17:54:31
Job time : 116 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: September 9, 2004, 17:52:38 ; Search time 127 Seconds
(without alignments)
55.553 Million cell updates/sec

Title: US-09-509-482D-21

Perfect score: 118

Sequence: 1 MEWRNKRSDMLSMVLRAGVE 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgm2_6/ptodata/2/pubaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgm2_6/ptodata/2/pubaa/US05_NEW_PUB.pep.*
- 4: /cgm2_6/ptodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgm2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgm2_6/ptodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgm2_6/ptodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgm2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgm2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgm2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgm2_6/ptodata/2/pubaa/US09_NEW_PUB.pep.*
- 13: /cgm2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgm2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgm2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgm2_6/ptodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgm2_6/ptodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgm2_6/ptodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	43.2	574	12	US-10-425-114-43706 Sequence 43706, A
2	51	43.2	584	12	US-10-425-114-70218 Sequence 70218, A
3	51	43.2	715	12	US-10-425-114-45401 Sequence 45401, A
4	51	43.2	923	12	US-10-425-114-44005 Sequence 44005, A
5	51	43.2	960	16	US-10-767-701-46165 Sequence 46165, A
6	51	43.2	968	12	US-10-425-114-38020 Sequence 38020, A
7	51	43.2	968	16	US-10-437-863-144756 Sequence 144756, A
8	51	43.2	990	16	US-10-437-863-144757 Sequence 144757, A
9	48	40.7	213	12	US-10-282-122A-72274 Sequence 72274, A
10	47	39.8	698	15	US-10-369-493-10484 Sequence 10484, A
11	46.5	39.4	638	12	US-10-282-122A-78426 Sequence 78426, A
12	46.5	39.4	731	12	US-10-282-122A-60220 Sequence 60220, A
13	46	39.0	102	15	US-10-284-049-3525 Sequence 3525, Ap
14	46	39.0	300	15	US-10-369-493-9779 Sequence 9779, Ap
15	46	39.0	303	12	US-10-424-599-229583 Sequence 229583, A

16	46	39.0	321	14	US-10-080-170-185	Sequence 185, App
17	46	39.0	321	16	US-10-080-170-185	Sequence 185, App
18	46	39.0	324	14	US-10-080-170-566	Sequence 566, App
19	46	39.0	324	16	US-10-080-170-566	Sequence 566, App
20	46	39.0	363	14	US-10-081-816-51	Sequence 81, Appl
21	45.5	38.6	349	15	US-10-369-493-16589	Sequence 16589, A
22	45	38.1	94	12	US-10-424-599-240589	Sequence 240589, A
23	45	38.1	107	12	US-10-424-599-196073	Sequence 196073, A
24	45	38.1	116	16	US-10-437-963-124041	Sequence 124041, A
25	45	38.1	609	12	US-10-425-114-43306	Sequence 43306, A
26	45	38.1	643	12	US-10-425-114-59001	Sequence 59001, A
27	44.5	37.7	697	16	US-10-437-963-160359	Sequence 160359, A
28	44	37.3	103	12	US-10-424-599-156449	Sequence 156449, A
29	44	37.3	233	12	US-10-424-599-171601	Sequence 171601, A
30	44	37.3	522	14	US-10-369-493-967	Sequence 967, App
31	44	37.3	556	15	US-10-437-963-146343	Sequence 146343, A
32	44	37.3	751	16	US-09-757-049A-1	Sequence 1, Appli
33	44	37.3	802	9	US-10-153-273-8	Sequence 8, Appli
34	44	37.3	921	13	US-10-437-963-154999	Sequence 154999, A
35	44	37.3	972	16	US-10-425-114-47259	Sequence 47259, A
36	44	37.3	989	12	US-09-864-408A-3988	Sequence 3988, Ap
37	43	36.4	58	11	US-10-437-963-104703	Sequence 104703, A
38	43	36.4	86	16	US-10-424-599-276925	Sequence 276925, A
39	43	36.4	159	12	US-10-362-327-34	Sequence 34, Appl
40	43	36.4	321	12	US-10-156-761-14331	Sequence 14331, A
41	43	36.4	373	14	US-10-282-122A-72287	Sequence 72287, A
42	43	36.4	409	12	US-10-425-114-40141	Sequence 40141, A
43	43	36.4	673	12	US-10-282-122A-75847	Sequence 75847, A
44	43	36.4	740	12	US-09-815-342-14001	Sequence 14001, A
45	43	36.4	751	9		

ALIGNMENTS

RESULT 1

US-10-425-114-43706
; Sequence 43706, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43706
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700571335_FLI.pep
US-10-425-114-43706

Query Match 43.2%; Score 51; DB 12; Length 574;
Best Local Similarity 56.2%; Pred. No. 46;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Cy 2 EWRNKRSDMLSMVLR 17

Db 84 EWRNKRSDMLSMVLR 99

RESULT 2

US-10-425-114-70218
; Sequence 70218, Application US/10425114
; Publication No. US20040034888A1

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 70218
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLMOL7168D04_FLI.pep
US-10-425-114-70218

Query Match          43.2%; Score 51; DB 12; Length 584;
Best Local Similarity 56.2%; Pred. No. 47;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      2 EWRNKRSDWLSMVL 17
      || : || || ||
Db      94 EWSEKRDWLLSEL 109

RESULT 3
US-10-425-114-45401
; Sequence 45401, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45401
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700152625_FLI.pep
US-10-425-114-45401

Query Match          43.2%; Score 51; DB 12; Length 715;
Best Local Similarity 56.2%; Pred. No. 57;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      2 EWRNKRSDWLSMVL 17
      || : || || ||
Db      225 EWSEKRDWLLSEL 240

RESULT 4
US-10-425-114-44005
; Sequence 44005, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44005
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700349524_FLI.pep
US-10-425-114-44005

Query Match          43.2%; Score 51; DB 12; Length 923;
Best Local Similarity 56.2%; Pred. No. 73;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      2 EWRNKRSDWLSMVL 17
      || : || || ||
Db      433 EWSEKRDWLLSEL 448

RESULT 5
US-10-767-701-46165
; Sequence 46165, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46165
; LENGTH: 960
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C86_1.pep
US-10-767-701-46165

Query Match          43.2%; Score 51; DB 16; Length 960;
Best Local Similarity 56.2%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      2 EWRNKRSDWLSMVL 17
      || : || || ||
Db      470 EWSEKRDWLLSEL 485

RESULT 6
US-10-425-114-38020
; Sequence 38020, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
```

```
; SEQ ID NO 38020
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700154435_FLI.pep
US-10-425-114-38020

Query Match      43.2%; Score 51; DB 12; Length 968;
Best Local Similarity 56.2%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy  2 EWRNKRSDWLSMVL 17
    || :||| |||
Db  478 EWSEKRDWLSSEL 493

RESULT 7
US-10-437-963-144756
; Sequence 144756, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144756
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45541C.1.pep
US-10-437-963-144756

Query Match      43.2%; Score 51; DB 16; Length 968;
Best Local Similarity 56.2%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy  2 EWRNKRSDWLSMVL 17
    || :||| |||
Db  478 EWSEKRDWLSSEL 493

RESULT 8
US-10-437-963-144757
; Sequence 144757, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 144757
```

```
; LENGTH: 990
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_45542C.1.pep
US-10-437-963-144757

Query Match      43.2%; Score 51; DB 16; Length 990;
Best Local Similarity 56.2%; Pred. No. 78;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy  2 EWRNKRSDWLSMVL 17
    || :||| |||
Db  500 EWSEKRDWLSSEL 515

RESULT 9
US-10-282-122A-72274
; Sequence 72274, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Orlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72274
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (17)..(17)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-72274

Query Match      40.7%; Score 48; DB 12; Length 213;
Best Local Similarity 47.4%; Pred. No. 48;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 1;
```

QY 1 MEWRNKRSDWLSMVRTA 19
DB 132 LSWTNKK--DWLRLVLIATA 148

RESULT 10
US-10-369-493-10484

; Sequence 10484, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10484
; LENGTH: 598
; TYPE: PRT
; ORGANISM: *Sphingomonas aromaticivorans*
US-10-369-493-10484

Query Match 39.8%; Score 47; DB 15; Length 698;
Best Local Similarity 36.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 WRNKRSDWLSMVRTAGV 21
DB 660 WNAERSQDWYDFITTAGV 678

RESULT 11
US-10-282-122A-78426

; Sequence 78426, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60220
; LENGTH: 731
; TYPE: PRT
; ORGANISM: *Klebsiella pneumoniae*

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78426
; LENGTH: 638
; TYPE: PRT
; ORGANISM: *Yersinia pestis*
US-10-282-122A-78426

Query Match 39.4%; Score 46.5; DB 12; Length 638;
Best Local Similarity 52.9%; Pred. No. 2.3e+02;
Matches 9; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 5 NKRSWLSMVRTAGV 21
DB 59 NKXR-EWLPMLMQTLGV 74

RESULT 12
US-10-282-122A-60220

; Sequence 60220, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60220
; LENGTH: 731
; TYPE: PRT
; ORGANISM: *Klebsiella pneumoniae*

```

Query Match      39.0%; Score 46; DB 15; Length 300;
Best Local Similarity 45.0%; Pred.No. 1.3e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      3  WRNKRSDWLSMVLETAGVE 22
      ||| : ||| ||| |||
Db      152 WRKLOESDMAGIVLAAGVK 171

```


<hr/>					
; ORGANISM: Plasmodium falciparum					
US-08-568-459A-8					
Query Match 37.3%; Score 44; DB 2; Length 921;					
Best Local Similarity 60.0%; Pred. No. 2e+02;					
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;					
Qy	2 EWNKRSW 11				
Db	206 EWINQRTW 215				
RESULT 9					
US-08-487-826B-8					
; Sequence 8, Application US/08487826B					
; Patent No. 5993827					
; GENERAL INFORMATION:					
; APPLICANT: Sim, Kim I.					
; APPLICANT: Chitnis, Chetan					
; APPLICANT: Miller, Louis H.					
; APPLICANT: Peterson, David S.					
; APPLICANT: Su, Xin-zhaun					
; APPLICANT: Wellens, Thomas E.					
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX					
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS					
; NUMBER OF SEQUENCES: 45					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Knobbe Martens Olson & Bear					
; STREET: 620 Newport Center Drive 16th Floor					
; CITY: Newport Beach					
; STATE: California					
; COUNTRY: US					
; ZIP: 92660					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: PatentIn Release #1.0, Version #1.25					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/487,826B					
; FILING DATE: 10-SEP-1993					
; CLASSIFICATION: 435					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Israelson, Ned					
; REGISTRATION NUMBER: 29,655					
; REFERENCE/DOCKET NUMBER: NIH121.001CP1					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (619) 235-8550					
; TELEFAX: (619) 235-0176					
; INFORMATION FOR SEQ ID NO: 8:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 921 amino acids					
; TYPE: amino acid					
; STRANDEDNESS: single					
; TOPOLOGY: linear					
; MOLECULE TYPE: protein					
; HYPOTHEICAL: NO					
; ORIGINAL SOURCE:					
; ORGANISM: Plasmodium falciparum					
US-08-487-826B-8					
Query Match 37.3%; Score 44; DB 2; Length 921;					
Best Local Similarity 60.0%; Pred. No. 2e+02;					
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;					
Qy	2 EWNKRSW 11				
Db	206 EWINQRTW 215				
RESULT 10					
US-09-210-288-8					
; Sequence 8, Application US/09210288					
; GENERAL INFORMATION:					
; APPLICANT: Bernstein, Harold S.					
; APPLICANT: Coughlin, Shaun R.					
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle					
; FILE REFERENCE: UCSF-020/010S					
; CURRENT APPLICATION NUMBER: US/09/156,316					
; CURRENT FILING DATE: 1998-09-18					
; EARLIER APPLICATION NUMBER: 60/060,688					
; EARLIER FILING DATE: 1997-09-22					
; NUMBER OF SEQ ID NOS: 12					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 1					
; LENGTH: 802					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-156-316-1					
Query Match 37.3%; Score 44; DB 3; Length 802;					
Best Local Similarity 50.0%; Pred. No. 1.8e+02;					
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;					
Qy	4 RNKRSWLMLVLTAGV 21				
Db	268 RKKSDDLPSAILQTSGV 285				
RESULT 8					
US-08-568-459A-8					
; Sequence 8, Application US/08568459A					
; Patent No. 5849306					
; GENERAL INFORMATION:					
; APPLICANT: Sim, Kim I.					
; APPLICANT: Chitnis, Chetan					
; APPLICANT: Miller, Louis H.					
; APPLICANT: Peterson, David S.					
; APPLICANT: Su, Xin-zhaun					
; APPLICANT: Wellens, Thomas E.					
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX					
; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS					
; NUMBER OF SEQUENCES: 37					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Knobbe Martens Olson & Bear					
; STREET: 620 Newport Center Drive 16th Floor					
; CITY: Newport Beach					
; STATE: California					
; COUNTRY: US					
; ZIP: 92660					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: PatentIn Release #1.0, Version #1.25					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/568,459A					
; FILING DATE: 07-DEC-1995					
; CLASSIFICATION: 435					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Israelson, Ned					
; REGISTRATION NUMBER: 29,655					
; REFERENCE/DOCKET NUMBER: NIH121.001CP1					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (619) 235-8550					
; TELEFAX: (619) 235-0176					
; INFORMATION FOR SEQ ID NO: 8:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 921 amino acids					
; TYPE: amino acid					
; STRANDEDNESS: single					
; TOPOLOGY: linear					
; MOLECULE TYPE: protein					
; HYPOTHEICAL: NO					
; ORIGINAL SOURCE:					

	US-08-568-459A-8	ORGANISM: Plasmodium falciparum	
	Query Match	37.3%; Score 44; DB 2; Length 921;	
	Best Local Similarity	60.0%; Pred. No. 2e+02;	
	Matches	6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
Qy	2 EWNKRSRW 11		
Dd	206 EWINQRTEW 215		
	RESULT 9		
	US-08-487-826B-8		
	; Sequence 8, Application US/08487826B		
	; Patent No. 5993827		
	; GENERAL INFORMATION:		
	; APPLICANT: Sim, Kim I.		
	; APPLICANT: Chitnis, Chetan		
	; APPLICANT: Miller, Louis H.		
	; APPLICANT: Peterson, David S.		
	; APPLICANT: Su, Xin-zhaun		
	; APPLICANT: Wellens, Thomas E.		
	; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX		
	; AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS		
	; NUMBER OF SEQUENCES: 45		
	; CORRESPONDENCE ADDRESS:		
	; ADDRESSEE: Knobbe Martens Olson & Bear		
	; STREET: 620 Newport Center Drive 16th Floor		
	; CITY: Newport Beach		
	; STATE: California		
	; COUNTRY: US		
	; ZIP: 92660		
	; COMPUTER READABLE FORM:		
	; MEDIUM TYPE: Floppy disk		
	; COMPUTER: IBM PC compatible		
	; OPERATING SYSTEM: PC-DOS/MS-DOS		
	; SOFTWARE: Patent In Release #1.0, Version #1.25		
	; CURRENT APPLICATION DATA:		
	; APPLICATION NUMBER: US/08/487,826B		
	; FILING DATE: 10-SEP-1993		
	; CLASSIFICATION: 435		
	; ATTORNEY/AGENT INFORMATION:		
	; NAME: Israelson, Ned		
	; REGISTRATION NUMBER: 29,655		
	; REFERENCE/DOCKET NUMBER: NIH121.001CP1		
	; TELECOMMUNICATION INFORMATION:		
	; TELEPHONE: (619) 235-8550		
	; TELEFAX: (619) 235-0176		
	; INFORMATION FOR SEQ ID NO: 8:		
	; SEQUENCE CHARACTERISTICS:		
	; LENGTH: 921 amino acids		
	; TYPE: amino acid		
	; STRANDEDNESS: single		
	; TOPOLOGY: linear		
	; MOLECULE TYPE: protein		
	; HYPOTHEICAL: NO		
	; ORIGINAL SOURCE:		
	; ORGANISM: Plasmodium falciparum		
	; US-08-487-826B-8		
	Query Match	37.3%; Score 44; DB 2; Length 921;	
	Best Local Similarity	60.0%; Pred. No. 2e+02;	
	Matches	6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
Qy	2 EWNKRSRW 11		
Dd	206 EWINQRTEW 215		
	RESULT 10		
	US-09-210-288-8		
	; Sequence 8, Application US/09210288		

Patent No. 6392026
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 520 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-8
Query Match 37.3%; Score 44; DB 4; Length 921;
Best Local Similarity 60.0%; Pred. No. 2e+02; 1; Indels 0; Gaps 0;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 EWRNKRSDW 11
|||:|:|:
Db 206 EWINQRTEW 215
RESULT 11
US-09-134-000C-5220
Sequence 5220, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5220
LENGTH: 395
TYPE: PRT

ORGANISM: Enterococcus faecalis
US-09-134-000C-5220
Query Match 36.4%; Score 43; DB 4; Length 395;
Best Local Similarity 38.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 2 EWRNKRSDWLSVLRTAGVE 22
|||:|:|:
Db 282 KWRFIQRSDYFQMGENTSSE 302
RESULT 12
US-09-489-039A-10154
Sequence 10154, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10154
LENGTH: 475
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10154
Query Match 36.4%; Score 43; DB 4; Length 475;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 3 WRNKRSDWLSVLR 18
|||:|:|:
Db 149 WRNKRWDFFSHYART 164
RESULT 13
US-09-543-681A-6761
Sequence 6761, Application US/09543681A
Patent No. 6505709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRAB
DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6761
LENGTH: 567
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6761
Query Match 35.4%; Score 43; DB 4; Length 567;
Best Local Similarity 32.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 7; Mismatches 6; Indels 4; Gaps 1;
QY 2 EWRNKRSDWLSMV---LRTAGVE 22
|||:|:|:
Db 34 QMGAKTSEWLTQYSDYDIRFSGIE 58
RESULT 14
US-08-897-843A-1
Sequence 1, Application US/08897843A

; Patent No. 6514493
; GENERAL INFORMATION:
; APPLICANT: DeLeo, Albert B.; Loftus, Douglas; Appella, Ettore
; TITLE OF INVENTION: CDNA CLONE FOR MURINE TUMOR
; TITLE OF INVENTION: REJECTION ANTIGEN GP110 AND TUMOR PEPTIDE VACCINE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Diane R. Meyers
; STREET: 600 Grant Street, 42nd Floor
; CITY: Pittsburgh
; STATE: PA
; COUNTRY: USA
; ZIP: 15219
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,843A
; FILING DATE:
; CLASSIFICATION: 424
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 938 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: NO
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; DEVELOPMENTAL STAGE: Embryo
; TISSUE TYPE: Embryo
; CELL LINE: NIH 3T3
; US-08-897-843A-1

Query Match 36.4%; Score 43; DB 4; Length 938;
Best Local Similarity 50.0%; Pred.No. 2.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 6 KKPSDWLSWLVLTAGV 21
Db 6 RRRSWLSVLVLYGV 21

RESULT 15
US-09-606-312-2
; Sequence 2, Application US/09606312
; Patent No. 6599732
; GENERAL INFORMATION:
; APPLICANT: Rayapati, P. John
; APPLICANT: Crafton, Corey M.
; TITLE OF INVENTION: Regulation of Carbon Assimilation
; FILE REFERENCE: 1533.093001
; CURRENT APPLICATION NUMBER: US/09/606,312
; CURRENT FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 60/141,001
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Medicago sativa
; US-09-606-312-2

Query Match 36.4%; Score 43; DB 4; Length 966;
Best Local Similarity 54.5%; Pred.No. 3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EWRNKKPSDWL 12

Db 475 EWSEKQEWL 485

Search completed: September 9, 2004, 17:55:51
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 17:38:46 ; Search time 120 Seconds
(without alignments)

51.800 Million cell updates/sec

Title: US-09-509-482D-21

Perfect score: 118

Sequence: 1 MEWNRKRSWLSWLVRTAGVE 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	72	2	AA09037 Human vit
2	118	100.0	450	2	AA09036 Human vit
3	118	100.0	477	2	AA09035 Human vit
4	48	40.7	213	6	ABU44350 Protein e
5	48	40.7	906	6	ABM67675 Phototab
6	47.5	40.3	63	3	AG15249 Arabidops
7	46.5	39.4	638	6	ABU50502 Protein e
8	46.5	39.4	731	6	ABU32296 Protein e
9	46	39.0	102	5	ABP42393 Human ova
10	46	39.0	321	5	ABU05534 M. tuberc
11	46	39.0	324	5	ABU05915 M. tuberc
12	46	39.0	363	5	AAE30531 Fruit fly
13	46	39.0	480	6	ADA36590 Acinetoba
14	45.5	38.6	431	6	AAO23463 Partial S
15	45.5	38.6	864	6	AAO23458 Sterculia
16	45.5	38.6	864	6	AAO23467 Sterculia
17	45	38.1	28	2	AAW78392 Finger F2
18	45	38.1	299	2	AAW89198 Aspartate
19	45	38.1	369	4	ABG26954 Novel hum
20	45	38.1	935	1	AP70362 Sequence
21	44.5	37.7	857	4	AB50562 C. elegans
22	44	37.3	185	3	AAV56159 HDNP1 let
23	44	37.3	185	4	ABG15759 Novel hum
24	44	37.3	214	6	ADA35408 Acinetoba
25	44	37.3	244	6	ABP78444 N. gonorr

26	44	37.3	325	3	AA06659	Ag06659 Arabidops
27	44	37.3	362	3	AA06658	Ag06658 Arabidops
28	44	37.3	365	3	AA06657	Ag06657 Arabidops
29	44	37.3	449	6	ADA36194	Ada36194 Acinetoba
30	44	37.3	614	3	AA757116	AA757116 Neisseria
31	44	37.3	715	4	AA896249	AA896249 Putative
32	44	37.3	793	2	AA70234	AA70234 P. falcip
33	44	37.3	802	2	AA60570	AA60570 Hydroneph
34	44	37.3	802	4	AA61997	AA61997 Human CDC
35	44	37.3	802	4	AA635742	AA635742 Human CDC
36	44	37.3	802	6	ABU12121	ABU12121 Human CDC
37	44	37.3	827	6	ABO53072	ABO53072 Human put
38	44	37.3	921	2	AAW22480	AAW22480 Plasmodiu
39	44	37.3	921	2	AA77902	AA77902 P. falcip
40	44	37.3	970	2	AAW89409	AAW89409 Maize pho
41	44	37.3	1856	2	AA721805	AA721805 Reading f
42	44	37.3	1856	3	AA783274	AA783274 Polypepti
43	43.5	36.9	436	7	ADC01008	ADC01008 Enterohae
44	43	36.4	58	5	ABP33021	ABP33021 Human ORF
45	43	36.4	95	3	AA42244	AA42244 Human ORF

ALIGNMENTS

RESULT 1
AA09037
ID: AA09037 standard; protein; 72 AA.
XX
AC AA09037;
XX
XX
DT 06-JUL-1999 (first entry)
XX
DE Human vitamin D receptor (VDR) gene transcript 10 amino acid sequence.
XX
XX Vitamin D receptor; VDR; hVDR; variant; isoform; SRC-1; GRIP-1; TFLIB;
XX cofactor; human.
XX
XX Homo sapiens.
OS
PN WO9916872-A1.
XX
PD 08-APR-1999.
XX
PF 29-SEP-1998; 98WO-AU000817.
XX
PR 29-SEP-1997; 97AU-00009500.
XX
XX (GARV-) GARVAN INST MEDICAL RES.
XX
XX Crofts LA, Hancock MS, Morrison NA, Eisman JA;
XX
XX WPI; 1999-263693/22.
XX
XX N-PSDB; AAX34790.
XX
XX New polynucleotides which encode novel isoforms of the human vitamin D
XX receptor or variant transcripts for hVDR.
XX
XX Disclosure; Fig 7; 56pp; English.

IN

The invention relates to isolated polynucleotides which encode novel isoforms of the human vitamin D receptor (VDR) or variant transcripts for hVDR. The polynucleotides are useful in methods for detecting agonist and/or antagonist compound of a VDR isoform. An increase or decrease in activity of the receptor may be detected by measuring changes in interactions with known cofactors (e.g. SRC-1, GRIP-1 and TFLIB) or unknown cofactors (e.g. through use of the dual hybrid system). The polynucleotides shown in AAX34797, AAX34791 and AAX34792 (corresponding to hVDR exon sequences) may be useful as probes for the detection of VDR variant transcripts and for assessing cell or tissue-specific expression of variant transcripts. The present sequence represents the amino acid sequence of hVDR gene transcript 10

RESULT 3
AAV09035
ID AAV09035 standard: protein: 477 AA.

XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00348933.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA48220.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 72274; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

XX DE Photorhabdus luminescens protein sequence #772.
XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
XX KW detection; food; gene expression; plant; animal; microorganism; toxin;
XX KW antibiotic; biopesticide; virulence factor; disease model; plague;
XX KW whooping cough.
XX OS Photorhabdus luminescens.
XX PN WO200294867-A2.
XX PD 28-NOV-2002.
XX PF 07-FEB-2002; 2002WO-IB003040.
XX PR 07-FEB-2001; 2001FR-00001659.
XX PA (INSP) INST PASTEUR.
XX PA (CNRS) CNRS CENT NAT RECH SCI.
XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunat F, Danchin A;
XX PI Buchrieser C;
XX WPI; 2003-148459/14.
XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX PS Claim 2; SEQ ID NO 772; 1205pp; French.
XX CC The invention relates to the isolation of genes and their encoded
XX CC proteins from Photorhabdus luminescens. The isolated sequences are
XX CC sources of probes and primers for detecting the genome of *P. luminescens*
XX CC and related species; to study polymorphisms; for gene analysis and for
XX CC detection/amplification of the genes. Antibodies (Ab) raised against the
XX CC polypeptides encoded by the genes are used for detection/identification
XX CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that
XX CC carry a gene-containing vector are used to select compounds that
XX CC modulate, regulate, induce or inhibit expression of the genes in plants,
XX CC animals or microorganisms other than *P. luminescens* and are able to alter
XX CC response or sensitivity to toxins and antibiotics produced by *P.*
XX CC *luminescens*. Cells transformed to express the genes are useful for
XX CC recombinant production of the proteins, particularly toxins and
XX CC antibacterials useful as insecticides, bactericides and fungicides. The
XX CC genes, proteins, vectors containing the genes and Ab are also useful
XX CC therapeutically (to treat microbial infection by bacteria or fungi that
XX CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as
XX CC biopesticides. Other uses of the genes and the proteins are as virulence
XX CC factors and for identifying targets of human diseases for which *P.*
XX CC *luminescens* is a model (particularly plague and whooping cough). This
XX CC sequence represents one of the isolated *P. luminescens* proteins
XX SQ Sequence 906 AA;
Query Match 40.7%; Score 48; DB 6; Length 906;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 EWRNKGSDWLS 13
DB 699 EWRQPCSDWLS 710
RESULT 6
AAG15249
ID AAG15249 standard; protein; 63 AA.
XX AC AAG15249;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 15428.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 28-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134768P.

PR 20-MAY-1999; 99US-0134941P.

PR 21-MAY-1999; 99US-0135124P.

PR 24-MAY-1999; 99US-0135353P.

PR 25-MAY-1999; 99US-0135629P.

PR 27-MAY-1999; 99US-0136021P.

PR 28-MAY-1999; 99US-0136332P.

PR 01-JUN-1999; 99US-0136782P.

PR 03-JUN-1999; 99US-0137222P.

PR 04-JUN-1999; 99US-0137528P.

PR 07-JUN-1999; 99US-0137502P.

PR 08-JUN-1999; 99US-0137724P.

PR 10-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.

PR 14-JUN-1999; 99US-0138847P.

PR 16-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139453P.

PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139464P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140635P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140931P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143634P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144613P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 28-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 05-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.

PR 18-AUG-1999; 99US-0149426P.

PR 20-AUG-1999; 99US-0149722P.

PR 20-AUG-1999; 99US-0149723P.

PR 20-AUG-1999; 99US-0149939P.

PR 23-AUG-1999; 99US-0149902P.

PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.

PR 27-AUG-1999; 99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.

PR 27-AUG-1999; 99US-0151080P.

PR 30-AUG-1999; 99US-0151303P.

19-JUN-2003 (first entry)

Protein encoded by Prokaryotic essential gene #17823.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Klebsiella pneumoniae. --

W0200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0142923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohisen KL, Zyskind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH; WPI: 2003-029926/02.

N-PSDB; ACA36166.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 60220; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 8213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

Sequence 731 AA;

Query Match 39.4%; SCORE 46.5; DB 6; Length 731;

Best Local Similarity 40.9%; Pred. No. 2.5e+02;

Matches 9; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

2 EW---RNKRSDWLSWLVRTAG 20

CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 102 AA;

Query Match 39.0%; Score 46; DB 5; Length 102;
 Best Local Similarity 54.5%; Pred. No. 33;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MEWENKKSOW 11
 |||||:|:|
 Db 18 MNWRRRKXOW 28

RESULT 10
 ABU05534
 ID ABU05534 standard; protein; 321 AA.

XX AC ABU05534;

XX DT 08-APR-2003 (first entry)

XX DE M. tuberculosis and M. leprae marker protein #185.

XX XW Mycobacterioses; survival; virulence; protective antigen; vaccine;
 XW Mycobacterial disease; tuberculosis; leprosy.

XX OS Mycobacterium tuberculosis.
 OS Mycobacterium leprae.

XX PN WO200274903-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-IB001973.

XX PR 22-FEB-2001; 2001US-0270123P.

XX PA (INSP) INST PASTEUR.

XX PI Cole S;

XX DR WPI; 2002-759885/82.

XX PT Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. leprae.

XX PS Claim 17; Page 363-364; 874pp; English.

XX CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC leprae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a marker protein from
 CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention

XX SQ Sequence 321 AA;

Query Match 39.0%; Score 46; DB 5; Length 321;
 Best Local Similarity 53.3%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MEWENKKSOW 11
 |||||:|:|
 Db 18 MNWRRRKXOW 28

QY 2 EWRNKRSDWLSMVL 16
 |||||:|:|
 Db 267 EWRNKRSDWLSMVL 281

RESULT 11

ABU05915

ID ABU05915 standard; protein; 324 AA.

XX AC ABU05915;

XX DT 08-APR-2003 (first entry)

XX DE M. tuberculosis and M. leprae marker protein #566.

XX XW Mycobacterioses; survival; virulence; protective antigen; vaccine;
 XW Mycobacterial disease; tuberculosis; leprosy.

XX OS Mycobacterium tuberculosis.

XX OS Mycobacterium leprae.

XX PN WO200274903-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-IB001973.

XX PR 22-FEB-2001; 2001US-0270123P.

XX PA (INSP) INST PASTEUR.

XX PI Cole S;

XX DR WPI; 2002-759885/82.

XX PT Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. leprae.

XX PS Claim 17; Page 780; 874pp; English.

XX CC This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomic sequences and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC leprae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a marker protein from
 CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention

XX SQ Sequence 324 AA;

Query Match 39.0%; Score 46; DB 5; Length 324;
 Best Local Similarity 53.3%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EWRNKRSDWLSMVL 16
 |||||:|:|
 Db 270 EWRNKRSDWLSMVL 284

RESULT 12

ABU05915

ID ABU05915 standard; protein; 363 AA.

XX AC ABU05915;

```

XX DT 24-FEB-2003 (first entry)
XX DE Fruit fly gustatory receptor protein, Gr93F4.
XX KW Fruit fly; gustatory receptor; Gr; odorant receptor; Or; pesticide;
XX KW Gr93F4.
XX OS Drosophila melanogaster.
XX PN W0200268593-A2.
XX PD 06-SEP-2002.
XX PF 22-FEB-2002; 2002WO-US005414.
XX PR 23-FEB-2001; 2001US-0271319P.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Axel R, Scott K;
XX DR WPI; 2002-698668/75.
XX PT Novel nucleic acid encoding insect gustatory or odorant receptor protein
XX PT useful for identifying a compound which specifically binds to the
XX PT receptor for controlling a pest population in an area.
XX PS Claim 3; Page 232-234; 264pp; English.
XX CC The invention relates to an isolated nucleic acid encoding an insect
XX CC gustatory receptor (Gr) or odorant receptor (Or) protein which comprises
XX CC seven transmembrane domains and a C-terminal domain comprising
XX CC consecutive amino acids. The invention is useful for identifying a
XX CC compound which activates the insect receptor or inhibits the activity of
XX CC the insect receptor. The purified insect receptor protein is embedded in
XX CC a lipid bilayer. The invention is sprayed for combating ingestion of
XX CC crops by pest insects, combating disease-carrying insects in an area and
XX CC controlling a pest population in an area. The invention is useful for
XX CC detecting the presence of insect gustatory or odorant receptor and for
XX CC inhibiting the function of insect gustatory or odorant receptor or in
XX CC biological fluids isolated from them. The invention is also useful for
XX CC identifying or isolating other insect receptors and for combating pest
XX CC nuisances. The present sequence is fruit fly gustatory receptor (Gr)
XX CC protein
XX CC
XX SQ Sequence 363 AA;
    Query Match          39.0%; Score 46; DB 5; Length 363;
    Best Local Similarity 46.7%; Pred. No. 1.3e+02;
    Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 WNRKRSRWLSMWLR 17
DB 29 WNRSRRWKWSVTLR 43

RESULT 13
ADA36590
ID ADA36590 standard; protein; 480 AA.
XX AC
XX AC ADA36590;
XX DT 20-NOV-2003 (first entry)
XX DE Acinetobacter baumannii protein #3751.
XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX KW plant biocontrol agent.
XX OS Acinetobacter baumannii.
XX PN US6562958-B1.

XX DT 13-MAY-2003.
XX PD
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX PR WPI; 2003-576092/54.
XX DR N-PSDB; ADA32464.
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.
XX PS Example; SEQ ID NO 7877; 328pp; English.
XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for
XX CC plants. The present sequence represents the amino acid sequence of an A.
XX CC baumannii protein.
XX SQ Sequence 480 AA;
    Query Match          39.0%; Score 46; DB 6; Length 480;
    Best Local Similarity 47.1%; Pred. No. 1.8e+02;
    Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEWRKRSRWLSMWLR 17
DB 185 LDWRNKQYDTLQKLR 201

RESULT 14
AAO23463
ID AAO23463 standard; protein; 431 AA.
XX AC
XX AC AAO23463;
XX DT 06-NOV-2003 (first entry)
XX DE Partial Sterculia foetida CPA-FAS protein expressed in BL21 cells.
XX KW Plant; cyclopropane fatty acid synthase; CPA-FAS; lubricant additive;
XX KW cyclopropane fatty acid; CPA-FA; cyclopropane fatty acid; vegetable oil;
XX KW CPE-FA; cosmetic; hard fat; food industry; enzyme; BL21.
XX OS Sterculia foetida.
XX PN W02003060079-A2.
XX PD 24-JUL-2003.
XX PF 20-DEC-2002; 2002WO-US041250.
XX PR 21-DEC-2001; 2001US-0345152P.
XX PR 03-JUL-2002; 2002US-0393937P.
XX PA (UNMS ) UNIV MICHIGAN STATE.
XX PI Bao X, Ohlrogge JB, Pollard MR;
XX KW WPI; 2003-627382/59.
XX DR N-PSDB; AAL56894.

```

Search completed: September 9, 2004, 17:52:03
Job time : 124 secs